

Engineering Immunological Solutions for Pandemic-Level Threats

Thesis by
Luis Fernando Caldera

In Partial Fulfillment of the Requirements for the degree of
Doctor of Philosophy

The Caltech logo, featuring the word "Caltech" in a bold, orange, sans-serif font, centered within a light orange rectangular background.

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2026
(Defended April 6, 2026)

© 2026

Luis Fernando Caldera
ORCID: 0009-0005-3254-6563

ACKNOWLEDGEMENTS

My journey through graduate school would not have been possible without the mentorship, friendship, and steadfast support of so many people.

Over the past decade, I've been shaped by exceptional mentors at each stage of my scientific development. Thank you Mark Kozlowski and Weslee Glenn for giving me my foundation in Dave Tirrell's lab at Caltech. Your patience and genuine enthusiasm for teaching inspired me to pursue research through my undergraduate years. Devang Amin, thank you for your mentorship at UC Berkeley in Phil Messersmith's lab. Working with you crystallized my desire to pursuing a Ph.D. someday. Cathy Drennan, Steve Cohen, and Paty Feliciano, thank you for welcoming me into your lab at MIT and cultivating such an inclusive environment where I could continue to grow as a scientist. And to Pat Holder, Yavuz Dagdas, and Christine Huang at Genentech, thank you for showing me what research and development looked like beyond the ivory tower and for deepening my passion for scientific inquiry.

I've also been blessed with a wonderful community of friends who have sustained me through this journey. To my Sanctuary Berkeley family, Dave, Corynne, Elliott, and Caiden Escalante, James Kendrick, Natalie Guzikowski, Aqshems Nichols, Mayra Luna, Paige and Jason Patrick, Wesley and Jing Wu, Brandon Borders, Olivia Lee, Danielle Speller, and Adelene Miranda, thank you for teaching me the importance of joy, good coffee, and spiritual growth. You made my time at Berkeley so memorable, especially during the pandemic years. Here in Pasadena, my extended family at Missio Community Church, Len and Amy Tang, Raymond Zhang, Evelyn Shan, Sami Chang, Hannah Manetsch, Albert Wandui, Mike Dao, and Shaylee Ungos (to name a few), have walked with me through both the celebrations and struggles. Your presence, listening ears, and willingness to love my children have meant everything to me.

My graduate studies represented a significant shift from biomaterials toward immunology and protein design. I'm grateful to Pamela Bjorkman and Stephen Mayo for welcoming me into their labs and trusting me to make that leap. Alex Cohen, thank you for your early

mentorship and for helping me find my footing in the lab as I acclimated to the group. In my HIV work, I gained indispensable knowledge from my lab mates Jennifer Keefe, Harry Gristick, Edem Gavor, Nick Koranda, Andy DeLaitsch, Kim Dam, and Anthony West. You all supported me during the challenging early years of my project when I felt overwhelmed trying to catch up with decades of research in a field entirely new to me. Adrian Bahn, thank you for being a great partner in the HIV projects. To my incredible undergraduate mentee Natalie Lee, thank you for the countless hours you invested in our work. Your commitment and thoroughness kept our project moving forward, and your thoughtful questions consistently challenged me to think more deeply. I'm grateful to have mentored you and am excited to see what you accomplish in your graduate work at MIT. Finally, thank you to my committee members Justin Bois, Ellen Rothenberg, and the late David Baltimore for your constructive comments and feedback throughout the years. Your perspectives have been instrumental in shaping me into a better scientist.

To my parents, Milburga Guzman and Alejandro Caldera, I owe you a profound debt. Your sacrifices as immigrants made my achievements possible, and you never hesitated to provide for me or make me feel like I could accomplish anything. The mindset you helped me develop is a big reason I've continued to push forward in higher education. Alex, thank you for remaining close to me and for dropping what you're doing to show up for me, especially for big moments like my defense. To my in-laws, Blanca Moreno and Juan Ortiz, I'm equally grateful for your generosity with your time and energy, especially during our early Pasadena days and when Camila was born.

Nicolas and Camila, you have kept me grounded in what's really important in life. Coming home to you both was the highlight of each day and I'm so thankful to be your dad. Finally, I owe much of my success to my incredible wife, Saharai. You have stuck by me since high school and seen it all. I would not be where I am today without you by my side for half my life, and I'm so grateful and excited to see the life we build together as I wrap up my time at Caltech. Thank you for everything you do for me and our family.

ABSTRACT

Pandemics remain among the most serious threats to global public health, especially when viruses can efficiently spread and persist in human populations. Two major examples are the AIDS pandemic caused by HIV-1, which emerged in the early 1980s, and the COVID-19 pandemic caused by SARS-CoV-2, which began in late 2019. Despite their differences in timescale and transmission, HIV-1 and SARS-CoV-2 exhibit high antigenic diversity that drives an ongoing arms race between viral escape and the development of vaccines and therapeutics. This thesis presents three molecular engineering strategies to confront that problem. For SARS-CoV-2, we utilized computational tools to generate mosaic nanoparticle vaccines displaying engineered SARS-CoV-2 or selected sarbecovirus receptor-binding domains (RBDs) sharing conserved antigenic features to drive immunity toward eliciting cross-reactive responses. In naïve and pre-vaccinated mice, our lead candidate, mosaic-7_{COM}, elicited broader and more potent cross-reactive antibody responses compared to our prior mosaic-8b candidate. For HIV-1, we engineered a stabilized CD4-Ig reagent specific to HIV-1 envelope (Env) glycoprotein with greatly reduced off-target recognition of MHC class II. The final CD4 design had markedly increased thermostability (>20°C) and a nearly 50-fold improvement in mammalian expression compared to the wild-type construct, highlighting its potential as a therapeutic biologic. Additionally, we developed a yeast-display screening platform to isolate nanobody (V_{HH}) domains against the caldera, a conserved, glycan-free epitope exposed on Env upon engagement with CD4. This campaign yielded five V_{HH}-Fc candidates that bind receptor-bound Env and established a foundation for future biologic development targeting a previously inaccessible site on the virus.

PUBLISHED CONTENT AND CONTRIBUTIONS

Throughout my tenure at Caltech, I had the privilege of working alongside amazing collaborators and contributed to two published works that constitute chapters 2 and 3 (1, 2).

1. E. Wang*, A. A. Cohen*, **L. F. Caldera***, J. R. Keeffe, A. V. Rorick, Y. M. Adia, P. N. P. Gnanaprasag, P. J. Bjorkman, A. K. Chakraborty, Designed mosaic nanoparticles enhance cross-reactive immune responses in mice. *Cell* 188, 1036-1050.e11 (2025). <https://doi.org/10.1016/j.cell.2024.12.015>.
2. A. J. Bahn-Suh, **L. F. Caldera**, P. N. P. Gnanaprasag, J. R. Keeffe, M. S. Seaman, P. J. Bjorkman, S. L. Mayo, Stabilized gp120-specific CD4 for next-generation HIV-1 inhibitors. *bioRxiv* [Preprint] (2026). <https://doi.org/10.64898/2026.03.24.713825>.

** These authors contributed equally.*

Chapter 2 (1) presents our work on designed mosaic nanoparticles for vaccination against SARS-CoV-2 variants of concern (VOCs) and closely related sarbecoviruses with zoonotic potential, done in collaboration with Arup Chakraborty, Ph.D. at the Massachusetts Institute of Technology (MIT) and his then graduate student, Eric Wang, now a research scientist at Google DeepMind. Arup and Eric were responsible for all the computational work that resulted in optimizing the selection process for receptor-binding domain (RBD) selection for our mosaic nanoparticles, which included SARS-CoV-2 WA1 mutants and wild type sarbecovirus RBDs. Alexander Cohen, a postdoctoral fellow in our group, conceived of the project alongside Eric and the principal investigators. I was responsible for much of the experimental setup and validation of Eric's proposed nanoparticle designs. I performed site-directed mutagenesis to generate the ten SARS-CoV-2 WA1 mutants and subsequent cloning for said mutants as well as any selected sarbecovirus RBDs that were not already made in our group. The Protein Expression Center (PEC) expressed and purified RBDs on our behalf via His-Tag affinity chromatography. I further purified the RBDs by size-exclusion chromatography (SEC). To characterize the RBDs, I performed SDS-PAGE and enzyme-linked immunosorbent assay (ELISA) to validate expected molecular weight and desired

binding profiles (i.e., preferential binding to class 3 and class 4 epitopes), respectively. I conjugated RBDs onto mi3 protein-based nanoparticles (made by others in the group) and further purified (via SEC) to arrive at our designed RBD nanoparticles (RBD-NPs). Our vaccine candidates were tested in mice, and I was responsible for evaluating sera via ELISA to assess the binding profiles for elicited antibodies. Neutralization assays were performed by Priyanthi Gnanapragasam in our group. Deep mutational scanning (DMS) experiments were conducted by Alexander Cohen. Eric Wang wrote much of the manuscript. Alexander Cohen and I contributed to experimental results, discussion, and methods. I additionally made figures 3, 4, 6, and S4 and helped review and edit the manuscript.

Chapter 3 (2) presents our work on an engineered CD4 as an anti-HIV-1 biologic, a project led by Adrian Bahn, a fellow Caltech graduate student also co-advised by Pamela Björkman, Ph.D. and Stephen Mayo, Ph.D. Alongside the principal investigators and Adrian, I helped conceive the study. Adrian was responsible for all the computational work that resulted in the design of thermostable, gp120-specific soluble CD4. I performed site-directed mutagenesis to generate the initial ten thermostable mutants, as well as SEC-purify (after expression and His-Tag affinity purification by the PEC) and characterize via SDS-PAGE and ThermoFluor assay to assess desired molecular weights and apparent melting temperatures (i.e., thermostability), respectively. Adrian performed all subsequent experimental work with the exception of polyreactivity assays, performed by Jennifer Keefe, Ph.D., and neutralization assays, performed by Priyanthi Gnanapragasam in our group and Michael Seaman, Ph.D. at Harvard University. I made figures 1b and 2a and helped review and edit the manuscript.

TABLE OF CONTENTS

Acknowledgements	iii
Abstract.....	v
Published Content and Contributions	vi
Table of Contents	viii
Nomenclature	x
Chapter 1: Introduction	1
1.1 Pandemics and the need for broadly protective countermeasures	1
1.2 SARS-CoV-2 Biology and Pathogenesis	4
1.3 COVID-19 Vaccine Development and the Mosaic Nanoparticle Approach	7
1.4 HIV-1 Biology and Pathogenesis	12
1.5 Efforts Toward HIV-1 Prevention and Treatment	16
1.6 Targeting Invariant Domains of HIV-1	20
1.7 Scope of Work	24
1.8 References	25
Chapter 2: Designed mosaic nanoparticles enhance cross-reactive immune responses in mice.....	48
2.1 Summary	48
2.2 Introduction.....	49
2.3 Results	52
2.4 Discussion	75
2.5 Materials and Methods.....	79
2.6 Resource Availability.....	83
2.7 Acknowledgements.....	84
2.8 Author Contributions	85
2.9 Declaration of Interests	85
2.10 References	86
Chapter 3: Stabilized gp120-specific CD4 for next-generation HIV-1 inhibitors.....	98
3.1 Summary	98
3.2 Introduction.....	99
3.3 Results	101
3.4 Discussion	115
3.5 Materials and Methods.....	117
3.6 Acknowledgements.....	125
3.7 Author Contributions	125
3.8 Competing Interest Statement.....	125
3.9 Extended Data	125
3.10 References	136

Chapter 4: Towards discovering nanobodies that target invariant domains in HIV-1	145
4.1 Summary	145
4.2 Introduction	146
4.3 Results	148
4.4 Discussion	161
4.5 Materials and Methods.....	165
4.6 References	170
Chapter 5: Conclusions and Outlook	177
5.1 Conclusions	177
5.2 Outlook.....	179
5.3 References	181

NOMENCLATURE

Alexa Fluor 488, 647 (AF488, AF647). Fluorescent dyes for protein labeling.

Angiotensin-converting enzyme 2 (ACE2). Host cell receptor for SARS-CoV-2 entry.

Antibody-dependent cellular cytotoxicity (ADCC). Immune effector mechanism mediated by antibody Fc domain.

Antibody-dependent cellular phagocytosis (ADCP). Phagocytic clearance mechanism mediated by antibody Fc domain.

B41. HIV-1 clade B Env sequence used in this study.

BG505. HIV-1 clade A Env sequence commonly used in SOSIP constructs.

Biotin-streptavidin. High-affinity protein-protein interaction used for protein immobilization and labeling.

Broadly neutralizing antibody (bNAb). Antibody capable of neutralizing diverse HIV-1 strains.

Caldera. Conserved, glycan-free region exposed between gp120 protomers in open Env trimers.

CD4-binding site (CD4bs). Region of gp120 that interacts with CD4.

CD4-induced (CD4i). Epitopes or conformational states exposed only after CD4 engagement.

CDR1, CDR2, CDR3. Three CDR loops in each antibody chain; CDR3 is typically longest and most diverse.

Class 1 epitope. RBD antibody binding site that blocks ACE2 and is highly variable between strains.

Class 2 epitope. RBD antibody binding site that blocks ACE2 but is less accessible than class 1.

Class 3 epitope. RBD antibody binding site that does not block ACE2 but is highly conserved.

Class 4 epitope. Cryptic, highly conserved RBD epitope exposed only in certain conformations.

Cluster of differentiation 4 (CD4). Primary receptor for HIV-1 on CD4⁺ T helper cells and other immune cells.

Complementarity-determining region (CDR). Hypervariable loops in antibody; variable domains that contact antigen.

Conformational change. Alteration in protein shape, often triggered by ligand binding.

Cryogenic electron microscopy (Cryo-EM). High-resolution structural technique using frozen samples.

Cryo-EM reconstruction. Three-dimensional protein structure determined by electron microscopy.

Crystal structure. Atomic-resolution protein structure determined by X-ray crystallography.

Deep mutational scanning (DMS). High-throughput experimental method to measure effects of all possible single mutations in a protein region.

Deep sequencing. High-throughput DNA sequencing to analyze library diversity and enrichment.

Dissociation constant (K_D). Measure of binding affinity; lower values indicate less dissociation, i.e., stronger binding.

Disulfide-stabilized SOSIP (DS-SOSIP). SOSIP with additional disulfide bond (201C-433C) that prevents CD4-induced conformational changes.

Enhanced CD4-immunoglobulin (eCD4-Ig). Fusion protein combining CD4, coreceptor-mimetic peptide, and antibody Fc domain.

Envelope glycoprotein (Env). HIV-1 surface protein responsible for receptor binding and membrane fusion.

Enzyme-linked immunosorbent assay (ELISA). Plate-based assay to measure protein binding.

Epitope. Specific region of an antigen recognized by an antibody.

Expi293. Suspension-adapted HEK293 cell line for transient protein expression.

Fluorescence-activated cell sorting (FACS). Flow cytometry technique to sort cells based on fluorescent markers.

Fragment antigen-binding (Fab). Antibody fragment containing one complete antigen-binding site.

Fragment crystallizable (Fc). Constant region of antibody that mediates effector functions.

gp120-specific CD4 (gCD4). Engineered CD4 variant with improved stability and HIV-1 Env specificity.

glycoprotein 120 (gp120). Surface subunit of Env responsible for receptor binding.

glycoprotein 41 (gp41). Transmembrane subunit of Env responsible for membrane fusion.

Glycan shield. Dense array of *N*-linked carbohydrates that protect viral proteins from antibodies.

Half-maximal effective concentration (EC₅₀). Concentration required to achieve 50% of maximal response.

Half-maximal inhibitory concentration (IC₅₀). Concentration required to inhibit 50% of activity.

Hexahistidine tag (His-tag). Protein purification tag consisting of six histidine residues.

HIV-1 clades. Genetic subtypes of HIV-1; major clades include A, B, C representing global diversity.

Human immunodeficiency virus type 1 (HIV-1). Retrovirus causing AIDS; target of therapeutic interventions described in this thesis.

Khosta-2. Recently discovered sarbecovirus resistant to most COVID-19 vaccines.

Kilodalton (kDa). Unit of molecular mass; 1 kDa = 1000 g/mol.

M48U1. CD4-mimetic peptide with higher affinity to gp120 than CD4 ($K_D = 0.0154$ nM vs. 7.68 nM for CD4).

Magnetic-activated cell sorting (MACS). Technique to separate cells using magnetic beads.

Mass photometry. Single-molecule technique to measure molecular mass in solution.

Mean fluorescence intensity (MFI). Average fluorescence signal in flow cytometry.

mi3. 60-mer icosahedral protein nanoparticle scaffold used for vaccine display.

Mosaic-2_{COM}. Mosaic-2 computationally optimized. Nanoparticle vaccine displaying 2 computationally designed SARS-CoV-2 variant RBDs.

Mosaic-5_{COM}. Mosaic-5 computationally optimized. Nanoparticle vaccine displaying 5 computationally designed SARS-CoV-2 variant RBDs.

Mosaic-7_{COM}. Mosaic-7 computationally optimized. Nanoparticle vaccine displaying 7 zoonotic sarbecovirus RBDs.

Mosaic-8b. Previously developed nanoparticle displaying 8 empirically selected sarbecovirus RBDs.

Mosaic nanoparticle. Vaccine platform displaying RBDs from multiple coronavirus strains on a single nanoparticle scaffold.

***N*-hydroxysuccinimide ester (NHS ester).** Reactive chemistry for covalent protein labeling (e.g., biotinylation, AlexaFluor labeling)

Nanobody. Alternative term for V_{HH} ; small (~15 kDa) single-domain antibody derived from camelids.

Negative-stain electron microscopy (nsEM). EM technique using heavy metal stains for contrast.

Open/closed Env. Conformational states of Env trimer; closed is prefusion, open is CD4-bound.

Paratope. Antigen-binding region of an antibody (counterpart to epitope).

PGT145. Broadly neutralizing HIV-1 antibody used for Env purification via affinity chromatography.

Potential *N*-linked glycosylation site (PNGS). Sequence motif (NXS/T) where *N*-linked carbohydrates can be attached.

Protein A. Bacterial protein that binds antibody Fc domains; used for antibody purification.

Protein Data Bank (PDB). Public repository for protein structure coordinates.

Protomer. Individual subunit within a multimeric protein complex.

Pseudovirus. Virus-like particle displaying viral spike proteins for neutralization studies.

Quaternary epitope. Antibody binding site formed by multiple protein subunits.

RaTG13. Bat coronavirus closely related to SARS-CoV-2.

Receptor-binding domain (RBD). The region of the SARS-CoV-2 spike protein that directly interacts with the ACE2 receptor on host cells.

S1/S2 subunits. Functional domains of the spike protein; S1 contains the RBD, S2 mediates membrane fusion.

Sarbecovirus. SARS-like betacoronavirus. Subgenus of betacoronaviruses including SARS-CoV, SARS-CoV-2, and related bat coronaviruses.

SARS-CoV-2 BA.5. Omicron subvariant that became dominant in 2022.

SARS-CoV-2 Beta. Variant of concern (B.1.351) first identified in South Africa.

SARS-CoV-2 WA1. Original SARS-CoV-2 reference strain from Washington state.

SHC014. Bat coronavirus that was subject to gain-of-function research.

Single-chain variable fragment (scFv). Antibody fragment consisting of V_H and V_L domains connected by a flexible linker.

Size-exclusion chromatography (SEC). Protein purification technique that separates by molecular size.

Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Technique to separate proteins by molecular weight.

Soluble CD4 (sCD4). Recombinant form of CD4 domains 1-2 used to trigger Env opening *in vitro*.

SOSIP. Stabilized, soluble form of HIV-1 Env trimer used for structural and immunological studies.

Spike protein. Major surface glycoprotein of SARS-CoV-2 responsible for receptor binding and membrane fusion.

SpyCatcher-SpyTag. Protein conjugation system enabling covalent attachment of antigens to nanoparticle scaffolds.

Superose 6, Superdex 200. Size-exclusion chromatography columns used for protein purification.

Surface plasmon resonance (SPR). Real-time technique to measure protein-protein binding kinetics.

Trimer. Protein complex composed of three subunits (e.g., HIV-1 Env).

TZM-bl assay. Cell-based assay using reporter cells to measure HIV-1 neutralization.

Variable domain of heavy chain of heavy-chain antibody (V_{HH}). Single-domain antibody (nanobody) derived from camelids.

V_{HH}-Fc. Fusion protein combining nanobody with antibody Fc domain.

WIV1. Bat coronavirus capable of binding human ACE2.

Yeast display. Technique to display proteins on yeast surface for screening and selection.

ZM197M. HIV-1 clade C Env sequence used in this study.

Chapter 1

INTRODUCTION

1.1 Pandemics and the need for broadly protective countermeasures

Throughout human history, infectious diseases have repeatedly spread beyond local outbreaks to become global pandemics. Two examples in recent times are the acquired immunodeficiency syndrome (AIDS) pandemic, which emerged in the early 1980s, and the coronavirus disease 2019 (COVID-19) pandemic, which began in late 2019 (1–4). Both diseases are caused by enveloped ribonucleic acid (RNA) viruses that emerged from zoonotic spillover, or transmission from animals to humans (5, 6).

AIDS is caused by human immunodeficiency virus 1 (HIV-1), a highly immune-evasive virus that has circulated globally for over four decades. HIV-1 arose through zoonotic transmission from African apes, with the pandemic lineage (group M) originating from simian immunodeficiency viruses (SIVs) in chimpanzees in central Africa (7). Zoonotic spillover into humans took place in the early twentieth century, with the earliest recognized cases happening in 1981 with the Centers for Disease Control and Prevention (CDC) reporting rare pneumonia (PCP) and Kaposi's sarcoma in previously healthy young gay men (8). Between the early 1980s and mid-1990s, an HIV/AIDS diagnosis was largely considered a “death sentence” due to limited treatment options and high mortality rates. Although antiretroviral therapy (ART) transformed AIDS from a fatal disease into a manageable chronic condition for many since then, HIV-1 remains a major global health burden, with more than 40 million people currently living with the virus and approximately 1.3 million new infections occurring each year (9, 10).

COVID-19 is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), a virus defined by its rapid respiratory transmission (11). Like HIV-1, SARS-CoV-2 appears to have emerged through zoonotic spillover. The earliest recognized cases were reported in Wuhan, China in late 2019, with current evidence suggesting a natural zoonotic origin linked

to the wildlife trade (12). This is further supported by recent evolutionary analyses showing no detectable change in selection pressure immediately prior to the virus's emergence in humans, suggesting that extensive pre-zoonotic adaptation was not required for sustained human-to-human transmission (13). The virus spread rapidly, with more than 70,000 confirmed cases reported by February 2020 despite major public health efforts to identify cases, trace contacts, and limit transmission (14, 15). Within months the virus had spread to every continent, including Antarctica, ultimately resulting in over 7 million confirmed deaths worldwide as of April 2024, though this is likely an underestimate (16). Excess mortality analyses suggest the true global toll reached well into the tens of millions (17).

While operating over different timescales, both pandemics highlight the challenge of developing targeted interventions that can keep pace with viral evolution. For example, HIV-1 has one of the highest known biological mutation rates at about 4.1×10^{-3} mutations per base per cell (18). This mutation frequency, driven by error-prone reverse transcriptase, has allowed the virus to amass immense genetic diversity both across globally circulating clades and within individual patients over the course of infection. This diversity is further compounded by recombination events that generate new circulating recombinant forms (CRFs). While SARS-CoV-2, by contrast, has a proofreading mechanism to ensure high-fidelity replication, it has nonetheless been able to generate numerous variants of concern (VOCs) over the years that allow the virus to progressively evade immune responses elicited by earlier strains of the virus and/or “outdated” vaccines (19, 20). This is made possible by the massive scale of transmission that creates ample opportunity for replication and recombination events to occur. Occasionally, prolonged infections in immunocompromised individuals also create opportunities for novel variants with enhanced transmissibility and immune evasion to arise (21).

The differences in evolutionary dynamics result from the distinct infection timelines of the two viruses. HIV-1 establishes a chronic infection over decades that primarily enables host-level viral evolution. Upon infecting memory CD4⁺ T cells, HIV-1 integrates itself into the host genome via its integrase enzyme and establishes a proviral reservoir that can remain

latent for years (22, 23). This allows the virus to remain largely invisible to the immune system, making it difficult to eradicate with current therapeutic strategies (24, 25). The infection progresses from an acute viremia phase 2-4 weeks post-infection (characterized by high viral loads and flu-like symptoms) to a prolonged asymptomatic period of gradual CD4⁺ T cell decline over 10+ years that ultimately leads to AIDS, making patients susceptible to opportunistic infections (26). Over time, the HIV-1 envelope (Env) glycoprotein responsible for host cell infection acquires extraordinary sequence diversity within patients and across clades that allow it to outpace strain-specific immune responses (27, 28). A person living with HIV-1 (PLWH) can transmit the infection to an estimated 2-5 people ($R_0 \sim 2-5$) in high-transmission settings through shared bodily fluids (e.g., sexual contact, blood exposure, vertical transmission from mother to child). This drops to $R_0 < 1$ for patients on ART who keep viral loads at undetectable levels, essentially preventing transmission (29, 30). On the other hand, SARS-CoV-2 spreads through short, acute infections that drive rapid population-level evolution. As a respiratory virus, SARS-CoV-2 spreads much quicker ($R_0 \sim 2-4$ for the ancestral strain, $R_0 \sim 8-18$ for later Omicron VOCs) through droplets and aerosols, especially in dense urban environments (31–33). SARS-CoV-2 targets angiotensin-converting enzyme 2 (ACE2)-expressing cells in the upper and lower respiratory tract and disease severity ranges from asymptomatic infection to acute respiratory distress syndrome (ARDS), multiorgan failure, and death, with risk strongly stratified by age and comorbidities (34, 35). Over the course of several infections, the virus accumulates mutations in regions on its spike (S) protein that improve antibody escape and transmissibility. Ultimately, these distinct evolutionary timelines allow the viruses to grow into global pandemics.

Taken altogether, we see a compelling case to develop broadly protective strategies against pandemic-level threats. This includes vaccines that can elicit cross-reactive immunity targeting conserved viral epitopes, biologics that exploit the virus's essential host receptor interactions, and antibody-based therapeutics that target invariant domains that cannot mutate without the virus losing fitness. To this end, this thesis presents molecular engineering strategies in each of these areas aimed at these two viruses.

1.2 SARS-CoV-2 Biology and Pathogenesis

1.2.1 The spillover potential of sarbecoviruses

SARS-CoV-2 belongs to the sarbecovirus subgenus within the genus *Betacoronavirus* (within the family Coronaviridae), which includes large enveloped, positive-sense, single-stranded RNA viruses encoding for several structural and accessory proteins (36, 37). Sarbecoviruses are further classified into four major clades depending on their host receptor use and receptor-binding domain (RBD) sequences: clades 1a, 1b, 2, and 3 (38). SARS-CoV-1 and SARS-CoV-2 fall under clades 1a and 1b, respectively, and use human ACE2 as their target host receptor. Bat-derived viruses in clades 2 and 3, by contrast, have varying (often limited) ACE2 binding and share less RBD sequence identity to viruses in clades 1a and 1b (39). Despite their differences, viruses across these clades are known for their zoonotic potential, or their ability to transmit from animals to humans. For example, the 2002-2003 SARS epidemic was caused by SARS-CoV-1, a bat-associated coronavirus that crossed into humans and caused widespread respiratory disease (40). More recently, the COVID-19 pandemic resulted from the spillover of SARS-CoV-2. Additionally, the 2012 Middle East respiratory syndrome (MERS) was caused by yet another coronavirus thought to have originated in bats, MERS-CoV (within the subgenus merbecovirus), after spillover from camels (41). Subsequent surveillance efforts further identified more bat-derived sarbecoviruses with the capacity to use human ACE2 to cause infection, such as WIV1 (clade 1a, from Yunnan, China), Yun11 (clade 2, from Yunnan, China), and BtKY72 (clade 3, from Kenya), to name a few (42–44). Given the ongoing threat of future zoonotic events, there is a strong rationale for developing vaccines that can provide protection from SARS-CoV-2 and its variants in addition to closely related sarbecoviruses with spillover potential (45).

1.2.2 Cross-reactive epitopes on sarbecoviruses are promising targets

The spike (S) glycoprotein of SARS-CoV-2 is a trimeric fusion protein that mediates both ACE2 receptor binding and subsequent membrane fusion. As a result, it is the primary target for neutralizing antibodies and vaccine efforts (39, 46). Proteolytic cleavage of each S protomer occurs at the S1/S2 boundary by furin-like proteases followed by TMPRSS2 at the S2' site to generate the S1 and S2 subunits responsible for receptor binding and membrane

fusion, respectively (34, 47). For S1 specifically, the RBD directly engages ACE2 through its receptor-binding motif (RBM) to initiate the conformational changes required for viral entry (48, 49). It is worth noting that in the spike's prefusion state, the RBDs across all three protomers alternate between “up” (receptor-accessible) and “down” (receptor-inaccessible) conformations (39). For infection to take place, the RBDs must be positioned “up” (**Figure 1**).

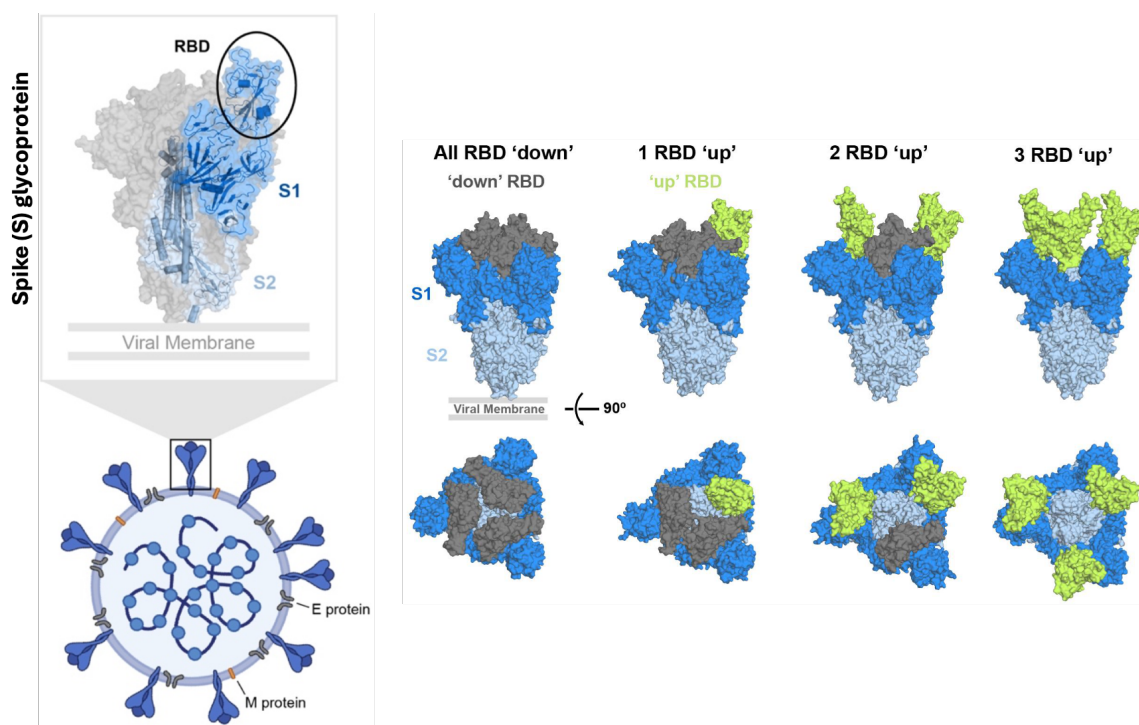


Figure 1. Structural organization and receptor-accessible conformation of the SARS-CoV-2 spike (s) glycoprotein. Adapted from Abernathy et al. (50)

Neutralizing antibodies targeting the SARS-CoV-2 RBD have been isolated, characterized, and categorized into four major epitope classes (class 1, 2, 3, and 4) based on where they bind on the RBD (51, 52). Class 1 antibodies bind the RBM in “up” RBDs and compete with ACE2 binding, neutralizing the virus by blocking viral entry. Class 2 antibodies similarly compete with ACE2 binding on the RBM, although they’re able to bind both “up” and “down” RBDs. Class 3 and 4 antibodies bind outside the ACE2 footprint and work through sterically hindering ACE2 binding, initiating Fc effector functions, or destabilizing the spike

protein. Class 3 antibodies can bind “up” or “down” RBDs, however class 4 antibodies target a region accessible only in “up” conformation. While class 1 and 2 antibodies are potent, they target more mutable regions and are thus susceptible to antibody escape in VOCs. For the same reason, they are not broadly neutralizing across related sarbecoviruses (i.e., not usually cross-reactive). Class 3 and 4 antibodies, however, target more conserved regions in the lower region of the RBD and have potential to be cross-reactive against sarbecoviruses in addition to SARS-CoV-2 VOCs. S309 (commercialized as sotrovimab, though no longer licensed for therapeutic use), for example, is a class 3 antibody isolated from a natural SARS-CoV-1 infection that neutralizes SARS-CoV-1, SARS-CoV-2, and other bat-derived sarbecoviruses like the clade 3 BM48-31 (51, 53). C118, a class 1/4 antibody whose epitope overlaps both class 1 and class 4 regions and competes with ACE2 binding, is also found to be cross-reactive due to engagement with conserved residues (54, 55). This makes the case for class 3/4- and class 1/4-targeting vaccine strategies.

1.2.3 SARS-CoV-2 VOCs evolved to escape class 1 and class 2 antibodies

SARS-CoV-2 has undergone extensive antigenic evolution since its emergence in late 2019, resulting in numerous VOCs over the following years. The first major one, Alpha (B.1.1.7), arose in late 2020 in the United Kingdom with the N501Y substitution in the RBM. This mutation greatly enhanced binding affinity and transmissibility (56). Additionally, this variant carried the D614G substitution on the S1 subunit, which increased infectivity by favoring the “up” RBD conformation. Beta (B.1.351), identified in South Africa shortly thereafter, introduced the K417N and E484K substitutions in the RBD that facilitated antibody escape against class 1 and class 2 antibodies, respectively, thereby reducing neutralization (57). Following Beta in 2021 was the Delta (B.1.617.2) variant, which combined possibly increased transmissibility (i.e., increased ACE2 binding) with enhanced immune evasion (e.g., electrostatically disrupting class 2 antibody binding) through its acquisition of the L452R and T478K substitutions on the RBM, respectively (58). By late November 2021, the Omicron (B.1.1.529) lineage took hold as the dominant circulating strain. The heavily mutated variant, along with its subvariants such as BA.1, BA.5, XBB.1.5, and JN.1, carried over 30 spike substitutions, including 15 in the RBD. This allowed Omicron

VOCs to escape existing therapeutic class 1 and class 2 monoclonal antibodies and weaken the effectiveness of first-generation vaccines (59, 60).

Deep mutational scanning (DMS) experiments on the SARS-CoV-2 RBD highlighted what we saw in the antigenic evolution of these VOCs: substitutions that improved immune evasion focused on class 1 and class 2 epitopes where binding of neutralizing antibodies could be ablated while simultaneously maintaining or enhancing ACE2 engagement (19, 61). By contrast, class 3 and class 4 epitopes, which were structurally and functionally more constrained, accumulated fewer escape substitutions. This insight further supported the idea that vaccines targeting these more conserved epitopes could provide more durable and broadly protective antibody responses compared to those focused on the more variable class 1 and class 2 epitopes.

While the first generation of COVID-19 vaccines, including monovalent and bivalent mRNA-LNP vaccines, provided significant protection against severe disease and death at the height of the pandemic, they progressively became less effective as the virus continued to evolve (62). The continued evolution toward immune evasion coupled with the ongoing threat of zoonotic spillover by related sarbecoviruses underscores the need for next-generation vaccine platforms capable of redirecting the immune response toward more conserved and cross-reactive epitopes.

1.3 COVID-19 Vaccine Development and the Mosaic Nanoparticle Approach

1.3.1 Limitations of current vaccine platforms

The rapid development and deployment of the mRNA-LNP vaccines against SARS-CoV-2 was a major biomedical breakthrough. The Pfizer-BioNTech BNT162b2 and Moderna mRNA-1273 vaccines, both encoding a prefusion-stabilized full-length spike protein, achieved >90% efficacy against symptomatic infection by the ancestral WA1 strain in phase 3 clinical trials and received emergency use authorization within 11 months of the viral genome's publication (63, 64). Following their release, there was a dramatic reduction in severe disease resulting in hospitalizations and death (65, 66).

As the virus continued to evolve, however, their effectiveness waned as the virus's mutations allowed it to evade the strain-specific responses elicited by the WA1-based vaccines. Against Omicron (BA.5), for example, neutralizing titers from two-dose vaccinations dropped 20-40-fold compared to WA1 (59, 60). Even boosters provided only marginal improvements against the evolving VOCs. To combat this, bivalent boosters incorporating both WA1 and Omicron BA.5 were developed (67). Unfortunately, clinical data indicated these bivalent boosters did not significantly outperform their monovalent counterparts with respect to eliciting cross-reactive neutralization against VOCs not included in the formulation (62). A potential complication was that of original antigenic sin (OAS), or immune imprinting, in which pre-vaccinated and/or previously infected patients, upon infection with a novel VOC, preferentially reactivated memory B cells against epitopes shared with the priming strain (i.e., WA1) rather than generating *de novo* responses against novel VOC-specific regions (68, 69). This bias could mean that updated vaccine boosters could be less effective than expected due to them expanding pre-existing, strain-specific responses rather than generating new cross-reactive ones. In the greater context of developing a universal pan-sarbecovirus vaccine, this imprinting poses a potential challenge as a majority of the global population has now been exposed to SARS-CoV-2 via infection and/or vaccination.

These limitations, coupled with the ongoing threat of zoonotic sarbecovirus spillover, motivate the development of next-generation pan-sarbecovirus vaccine strategies that target immune responses toward conserved regions across sarbecoviruses.

1.3.2 Multivalent antigen display vaccine platforms and the mosaic approach

Protein nanoparticle vaccines offer several immunological advantages over soluble protein antigens for eliciting cross-reactive immune responses. By presenting multiple copies of an antigen on a nanoparticle scaffold, these vaccines enable multivalent interactions with B cell receptors, increasing the effective binding strength through avidity (43, 70, 71). B cells whose receptors recognize epitopes that are repeated across many of the displayed antigens can engage several copies simultaneously, resulting in stronger activation signals. By

contrast, B cells that interact with only a single antigen copy experience weaker, monovalent engagement. This difference in binding strength can provide B cells with a competitive advantage during germinal center (GC) reactions, allowing them to expand and mature preferentially. This avidity-driven selection forms the basis of the mosaic nanoparticle vaccine approach.

A mosaic nanoparticle vaccine displays multiple distinct, yet related, antigens on a nanoparticle scaffold simultaneously. Unlike conventional nanoparticles that display only one type of antigen (homotypic), mosaic nanoparticles are engineered to induce broader, cross-reactive immune responses against multiple strains or variants of a virus, such as SARS-CoV-2 and other zoonotic sarbecoviruses. In our case, a nanoparticle displaying the RBDs of several distinct sarbecoviruses and/or SARS-CoV-2 VOCs engage B-cell receptors (BCRs) that can recognize epitopes conserved across adjacent RBDs at once. While individual BCRs can bind RBDs alone, their multivalent display on the nanoparticle permits cross-linking that drastically increases the overall avidity and subsequent activation of B cells recognizing shared epitopes. By increasing the number of antigens displayed on a mosaic nanoparticle, the likelihood of neighboring RBDs being distinct increases, facilitating this cross-reactive activation (43). By contrast, monovalent interactions with soluble protein antigens are susceptible to activation by immunodominant epitopes that may induce strain-specific responses.

As a proof-of-concept, a first-generation mosaic nanoparticle vaccine targeting sarbecoviruses was developed in our group by empirically selecting RBDs from SARS-CoV-2 and seven additional zoonotic sarbecoviruses based on their phylogenetic diversity across clades 1a, 1b, and 2 (43, 44). To enable flexible antigen display, the vaccine design leveraged the SpyTag-SpyCatcher protein conjugation system, which allows for the modular assembly of antigens onto self-assembling protein-based nanoparticle scaffolds (72). The SpyCatcher and SpyTag sequences form a spontaneous and irreversible isopeptide bond under mild reaction conditions, enabling antigens with a fused C-terminal SpyTag to be covalently coupled to a 60-mer nanoparticle scaffold (mi3) containing N-terminal SpyCatcher domains.

This modular conjugation system allows the same nanoparticle scaffold to display any SpyTagged antigen or combination of antigens, enabling us to synthesize both homotypic and mosaic display strategies with near-complete conjugation efficiency.

Using this strategy, pan-sarbecovirus mosaic-8b nanoparticles were constructed displaying RBDs from SARS-CoV-2, RaTG13, SHC014, Rs4081, pang17, RmYN02, Rf1, and WIV1. This vaccine successfully elicited cross-reactive antibody responses in mice and protected K18-hACE2 transgenic mice from lethal challenge against SARS-CoV-2 and SARS-CoV-1 (43, 44, 73). In nonhuman primates (NHPs), mosaic-8b similarly elicited cross-reactive responses and also protected from SARS-CoV-2 and SARS-CoV-1 challenges (44). In considering improvements to this vaccine candidate, we focused on optimizing the RBD selection process. While successful in its original design, the chosen eight RBDs might have been suboptimal due to the limitations of hand-selecting which viruses to choose. Viruses that are very closely related, such as clade 1a viruses WIV1 and SHC014, share variable immunodominant epitopes that might skew responses to be more strain-specific. Additionally, the now-dated SARS-CoV-2 Beta RBD chosen for this iteration could make the vaccine susceptible to imprinting in pre-vaccinated patients (73). To address this, computational approaches could be employed to efficiently optimize RBD selection in next-generation mosaic nanoparticle vaccines. Efforts to this end are the focus of Chapter 2 and described briefly in the following section.

1.3.3 Computationally optimizing RBD selection improves mosaic nanoparticles

Chapter 2 presents a computational framework for designing second-generation mosaic nanoparticle vaccines with RBDs that are specifically engineered or chosen to focus immune responses on conserved epitopes (74). The approach integrates three computational tools: (1) DMS escape data from yeast-displayed SARS-CoV-2 mutant libraries to identify residues whose mutation disrupts binding of class 1 and class 2 antibodies while preserving class 3 and class 4 binding, (2) a convoluted neural network trained on DMS expression data to predict whether multi-mutant RBD sequences will fold and express properly, and (3)

Aggrescan solubility scoring to ensure that designed RBDs do not aggregate (74–76). These computational studies were conducted in the Chakraborty lab at MIT.

For designed SARS-CoV-2 RBDs (i.e., WA1 RBD mutants), the pipeline began by generating 800,000 candidate six-mutation RBD sequences optimized to escape class 1 and class 2 antibodies (three mutations for each class). These were filtered down to 100,000 candidates predicted to express, then further filtered to 90,000 with favorable predicted solubility. Finally, mutants were ranked by total class 1/2 antibody escape to select the top 20,000 pairs. From these, sets of five RBD pairs (ten total RBDs) were selected for maximum mutational entropy to ensure diversity. For zoonotic sarbecovirus RBDs (without mutations), a complementary approach was used. 246 nonredundant sarbecovirus RBDs from the National Center for Biotechnology Information (NCBI) were aligned to the WA1 RBD and scored using a selective binding metric, selecting the top forty RBDs that maximized evasion of variable epitope antibodies while preserving conserved epitopes recognition. The final set of eight RBDs was chosen to minimize pairwise amino acid sequence identity.

Experimental validation in our group confirmed eight of ten designed SARS-CoV-2 WA1 mutants expressed well and bound class 3 and class 4 antibodies while ablating binding of class 1 and class 2 antibodies by enzyme-linked immunosorbent assay (ELISA), demonstrating that the neural network expression predictor was effective despite designing six-mutant RBDs (where 70% of single-point mutations alone eliminate expression). Five vaccine candidates were constructed with these data: mosaic-2a_{COM}, mosaic-2b_{COM}, mosaic-2c_{COM}, and mosaic-5_{COM} displaying designed two or five SARS-CoV-2 RBDs (for protection against circulating VOCs), and mosaic-7_{COM} with seven zoonotic sarbecovirus RBDs (excluding SARS-CoV-2 RBD; for pan-sarbecovirus protection).

These next-generation designs proved promising. Briefly, in immunized mice, mosaic-7_{COM} generated considerably enhanced binding antibody responses against all assessed RBDs compared to the original mosaic-8b and homotypic SARS-CoV-2 Beta nanoparticles, including sarbecovirus RBDs from clades 1a, 2, and 3. DMS profiling of mosaic-7_{COM}

antisera revealed preferential targeting of class 3 and class 4 RBD epitopes with reduced recognition of the class 1 and class 2 epitopes, confirming that the computational design strategy successfully redirected the antibody response from variable to conserved epitopes. In pre-vaccinated mice (i.e., three doses of mRNA-LNP prior to mosaic nanoparticle boosting), mosaic-7_{COM} elicited significantly higher neutralizing titers than mosaic-8b, mosaic-7 (mosaic-8b without SARS-CoV-2 RBD), and bivalent mRNA-LNP boosters against both zoonotic sarbecoviruses and highly-mutated Omicron VOCs, demonstrating its potential as a broad sarbecovirus vaccine for previously-infected and/or pre-vaccinated individuals.

1.4 HIV-1 Biology and Pathogenesis

1.4.1 HIV-1 structure and its limited Envelope glycoprotein coverage

HIV-1 is an enveloped RNA virus belonging to the genus *Lentivirus* within the family Retroviridae. Its approximately 9.7 kb single-stranded RNA genome encodes for several structural, regulatory, and accessory proteins (77). The mature virion is surrounded by a host-derived lipid membrane that displays the virus's only surface-accessible protein, the envelope (Env) spike glycoprotein (**Figure 2**). Env mediates viral entry by binding the CD4 receptor on the surface of CD4⁺ T cells and initiating membrane fusion. Each Env protomer is expressed as a gp160 precursor that is cleaved by furin to produce the gp120 and gp41 subunits responsible for facilitating receptor binding and membrane fusion, respectively (78, 79). Three gp120/gp41 heterodimers then self-assemble noncovalently to form trimeric spikes. The mature virion is decorated on its surface with approximately 7-14 Env spikes, a low density of spike coverage compared to other viruses (80, 81). This sparse coverage serves as an immune evasion strategy by making limiting inter-spike crosslinking (i.e., avidity) between bivalent antibodies and adjacent Env spikes (80).

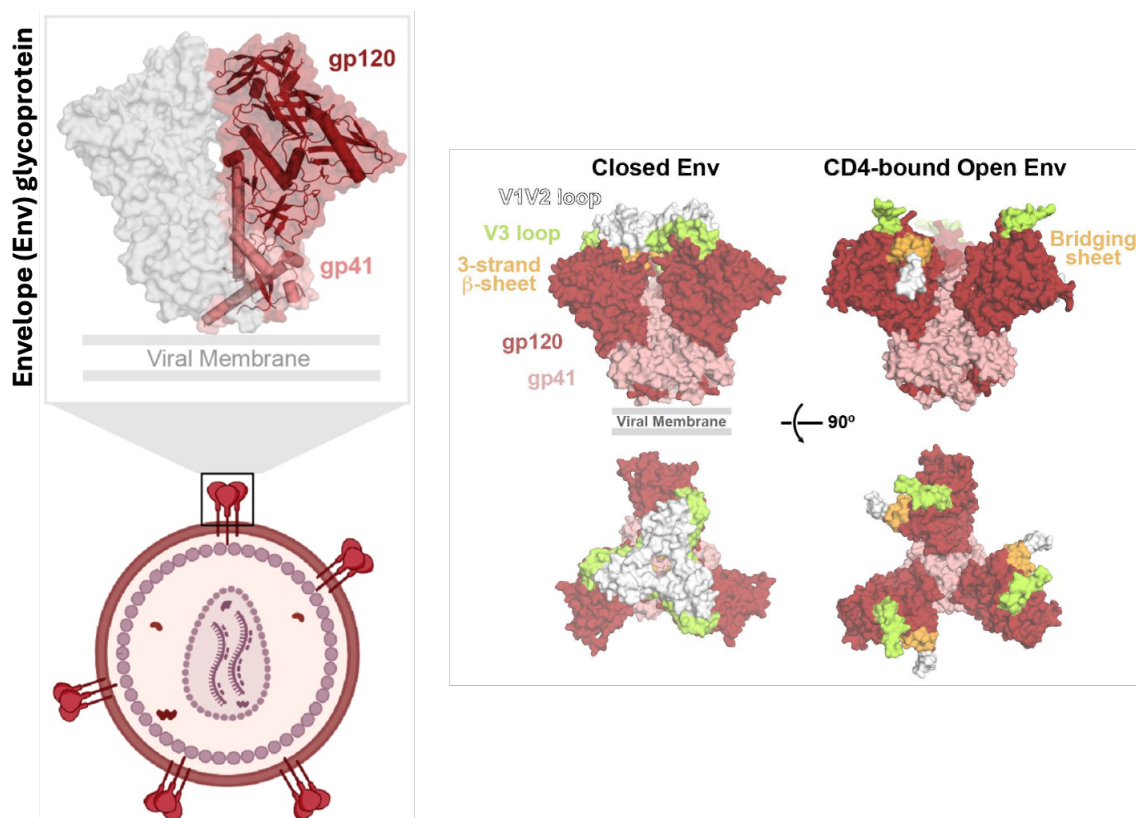


Figure 2. Structural organization and conformational rearrangement of the HIV-1 envelope (Env) glycoprotein. Adapted from Abernathy et al. (50)

In its “closed” prefusion conformation, Env adopts a compact structure in which variable loops at the apex of each protomer, termed V1V2, shield the co-receptor binding site and contribute to immune evasion (82). Additionally, the trimer is extensively and heterogeneously glycosylated, bearing a median of 94 *N*-linked glycans per trimer (83–85). These glycans, which are added by the host’s glycosylation machinery, cover much of the protein surface and sterically restrict antibody access to the underlying protein epitopes, essentially forming a “glycan shield” that masks the virus from the immune system. Upon engagement with CD4, Env undergoes a dramatic series of conformational changes, which have been elucidated through cryo-EM studies of Env in complex with soluble CD4 (sCD4) and CD4-mimetic compounds, such as M48U1 (86–90). First, gp120 protomers rotate outward by approximately 40 Å from the trimer axis, rearranging the V1V2 loops to expose the V3 region and co-receptor binding site (90). In this CD4-bound “open” conformation,

interior surfaces of the trimer completely buried in the “closed” state are revealed (**Figure 2**). This includes the inner surfaces of adjacent gp120 protomers and the top of the gp41 heptad repeat region. This conserved, glycan-free surface between the displaced gp120 protomers and atop gp41 constitute an invariant domain that our group has termed the “caldera” region due, in part, to its resemblance of the exposed basin of a volcanic crater.

1.4.2 The HIV-1 entry mechanism presents potential therapeutic targets

The CD4 receptor is expressed on the surface of immune cells, primarily on helper T cells, macrophages, and dendritic cells, and endogenously acts as a co-receptor with T-cell receptors (TCRs) to recognize antigens and initiate immune responses (91, 92). The CD4 binding site (CD4bs) on gp120 is a conserved cavity composed of discontinuous residues from the gp120 inner domain, outer domain, and bridging sheet. Despite being a promising target for broadly neutralizing antibodies, the CD4bs is partially shielded by the V1V2 and V3 loops and adjacent glycans in the “closed” trimer (78). HIV-1 entry occurs over a coordinated series of binding events initiated by CD4 binding the CD4bs (93, 94). gp120 rearrangements expose the V3 loop and bridging sheet that together form the co-receptor binding site. The co-receptor used depends on the HIV-1 strain. R5-tropic HIV-1 uses the C-C chemokine co-receptor CCR5, infecting macrophages and memory T cells. These strains dominate early infection because it is efficiently transmitted and replicates rapidly. X4-tropic HIV-1 uses the CXC-chemokine co-receptor CXCR4, targeting a broader range of T cells, and cause faster disease progression. Unfortunately, there also exist dual-tropic HIV-1, which can utilize either co-receptor (95–97). Co-receptor engagement triggers the release of the fusion peptide buried within gp41, which subsequently inserts itself into the target cell membrane. The gp41 N- and C-terminal heptad repeats then refold into a thermodynamically stable six-helix bundle (6HB). This coiled-coil structure brings the viral and host membranes together, ultimately driving membrane fusion until the viral capsid can enter the target cell cytoplasm (98, 99).

Each step of this cascade represents a potential therapeutic target, as evidenced by the successful implementation of ART cocktails with inhibitors for multiple different steps along

the HIV-1 lifecycle. For example, co-receptor engagement can be inhibited by the CCR5 antagonist maraviroc and membrane fusion can be blocked with the gp41-targeting peptide enfuvirtide (100, 101). CD4 binding can be blocked with CD4bs antibodies (e.g., VRC01) or CD4-based decoy biologics (e.g., CD4-Ig) (102, 103). CD4-induced “open” states can be exploited to target buried invariant regions for broadly neutralizing antibody discovery (e.g., caldera epitope). Later chapters in this thesis will focus on efforts toward the latter two.

1.4.3. HIV-1 antigenic diversity and immune evasion

HIV-1 exhibits the greatest antigenic diversity of any known human pathogen. The virus is classified into four groups (M, N, O, P) resulting from independent zoonotic transmissions from primates, with group M being responsible for the ongoing pandemic. Group M is further divided into subtypes (A through K, excluding E and I), or clades, as well as numerous circulating recombinant forms (CRFs). Clades C, B, and A account for the majority of infections in sub-Saharan Africa, the Americas and Europe, and East Africa, respectively (104–106). Inter-clade amino acid distances in Env can range from 25-35%, while diversity can exceed 15-20% within a single clade (107). This extraordinary diversity is driven by the combined results of HIV-1’s high mutation rate, rapid replication kinetics (approximately 10 to 10¹⁰ virions produced each day in untreated PLWH), frequent recombination events, and intense selective pressure from the host’s immune response (particularly from neutralizing antibodies and cytotoxic T cells) (18, 27, 108). Within a single PLWH, the virus continues evolving into a “swarm” of related but genetically distinct variants (quasispecies), with escape mutants emerging within weeks of the initial antibody response (28, 109).

Despite this seemingly endless variability, certain regions on Env are functionally constrained by what is required for receptor binding and membrane fusion. An obvious case is the preservation of the CD4bs on gp120, as mutations to this region that reduce CD4 binding directly impact viral fitness by reducing the virus’s ability to infect cells. Similarly, interior surfaces of the Env trimer exposed in the CD4-induced “open” state, such as the CD4-induced (CD4i) and caldera epitopes, are highly conserved across all HIV-1 clades and lack potential *N*-linked glycosylation sites (PNGSs) in a majority (>98%) of HIV-1 strains

(86, 87). This conservation reflects the functional requirement for these surfaces to preserve critical contacts for trimer formation (in “closed” Env) and CD4-induced conformational changes (in “open” Env). These constraints present promising targets for broadly protective therapeutics, as the virus cannot mutate these regions without incurring some loss in fitness.

1.5 Efforts Toward HIV-1 Prevention and Treatment

1.5.1 Challenges in developing an HIV-1 vaccine

Despite efforts to develop an effective HIV-1 vaccine over the past four decades, it remains a major unsolved challenge in biomedical science. Several obstacles, including the virus’s proviral latent reservoir, extraordinary antigenic diversity, and morphological characteristics (e.g., sparse spike coverage, glycan shield) are just a few that allow the virus to continually evade broad neutralization (109–111). To further complicate things, structure-based immunogen design using soluble versions of native Env trimers was not feasible for the first three decades due to their conformational instability. A breakthrough came in 2013, though, with the development of BG505 SOSIP.664, a strain A soluble Env (BG505) with stabilizing “SOSIP” mutations in its prefusion conformation (112). The name refers to a stabilized disulfide bond between gp120-gp41 subunits (SOS), an isoleucine-to-proline (IP) mutation (I559P) in the gp41 ectodomain, and a truncation at residue 664 to remove the insoluble transmembrane domain. While initially promising, these constructs, in addition to clade B (B41) and C (ZM197M, DU422) SOSIPs, were found to primarily elicit strain-specific, autologous neutralizing antibodies in immunization studies rather than broad cross-reactive ones necessary for effective widespread coverage (112–114).

Clinical trials have likewise produced limited success. To date, the RV144 Thai trial (conducted from 2003-2006) remains the only efficacy trial to demonstrate protection from HIV-1 acquisition (an approximately 31% reduction in infection risk) (115). The prime-boost strategy produced non-neutralizing antibodies primarily targeting the V1V2 region of gp120. Subsequent attempts to build on this success have fallen short, however, including the Uhambo (HVTN 702), Imbokodo (HVTN 705), and Mosaico (HVTN 706) trials conducted from 2016-2023, which ultimately demonstrated their vaccine candidates were safe but did

not provide significant protection against HIV-1 (*116, 117*). More recently, next-generation SOSIPs have been developed for more sophisticated germline-targeting (GT) approaches aiming to activate naïve B cells bearing germline precursors for known broadly neutralizing antibodies (bNAbs), specifically those targeting the CD4bs and V3 glycans (*118–122*). In this strategy, an initial priming immunogen is engineered to engage rare naïve B cells with receptors resembling the inferred germline precursors of a desired bNAb class. Sequential “shaping” and “polishing” boosts then slowly present more native-like Env features to guide antibody maturation toward more somatic hypermutation, affinity, neutralization potency, and breadth.

1.5.2 Passive immunotherapy with bNAbs falls short of universal protection

Despite shortcomings in vaccine development, researchers have nonetheless successfully been able to identify and isolate bNAbs from PLWH with excellent breadth and potency. These bNAbs recognize a limited number of epitopes on Env such as the CD4bs (e.g., VRC01, 3BNC117), the V1V2 region (e.g., PG9, PGT145), and the V3 region (e.g., 10-1074, 007), to name a few (*123–127*). In lieu of a vaccine, these bNAbs are useful as an alternative strategy for HIV-1 treatment. In early preclinical studies, passive transfer of 2F5, 4E10, and 2G12 bNAbs in nonhuman primates demonstrated potent protection against intravenous and mucosal simian-human immunodeficiency virus (SHIV) challenge (*128, 129*). VRC01-LS, a potent CD4bs bNAb with enhanced half-life Fc mutations (LS) also provided durable protection in macaques from repeated low-dose SHIV challenges (*130*).

In clinical trials, results have been encouraging, albeit with some limitations. Intravenous infusion of 3BNC117 suppressed plasma viremia for 28 days in PLWH during treatment interruption (*131*). However, some patients developed resistant viral strains while others showed viral rebound after antibody titers waned. 10-1074 showed similar viremia-suppressing activity but likewise resulted in the emergence of some viral resistance (*132*). Combination therapy using both 3BNC117 and 10-1074, targeting two distinct epitopes (CD4bs and a V3 glycan, respectively), provided sustained viral suppression for months after treatment interruption in a subset of patients with much less risk to developing viral

resistance, suggesting a multi-epitope targeting strategy can delay or prevent viral escape (133, 134). The AMP trial (Antibody Mediated Prevention; HVTN 704/HPTN 085) from 2016-2021 aimed to assess whether intravenous infusion of VRC01 could prevent HIV-1 infection. The trial administered VRC01 intravenously every eight weeks to high-risk individuals without HIV-1. While VRC01 prevented acquisition of VRC01-sensitive viral strains (concentration to inhibit virus by 80% (IC_{80}) $< 1 \mu\text{g/mL}$), it did not reduce overall HIV-1 acquisition rates compared to placebo because approximately 70% of circulating strains in the trial populations were already resistant to VRC01 (135). This result demonstrated that current antibody strategies targeting even highly conserved epitopes like the CD4bs were insufficient for universal protection, highlighting the need for therapeutics that target epitopes under the strongest possible functional constraints where escape mutations incur fitness costs too severe for the virus to tolerate.

1.5.3 CD4-based biologics exploit HIV-1's dependence on its host receptor

CD4-based biologics offer a different strategy for HIV-1 therapy, one that may help overcome some of the limitations faced by neutralizing antibodies. Unlike antibodies, CD4-based therapeutics exploit HIV-1's dependence on its primary host receptor. Because it is required for infection, mutations in gp120 that reduce binding to CD4-based therapeutics inherently also impair the virus's ability to bind cellular CD4 and infect target cells. In this way, CD4-based therapeutics create a direct link between immune escape and viral fitness cost (103, 136).

One of the earliest examples of this approach was CD4-Ig, an Fc-fusion protein composed of two N-terminal, immunoglobulin-like domains of human CD4 (D1 and D2) fused to the Fc region of human IgG, first described by Capon et al. in 1989 (137). The clinical candidate PRO-542 (CD4-IgG2) demonstrated antiviral activity in multiple clinical studies. In adult PLWH, a single intravenous dose produced a 0.5-log reduction in plasma viremia (138). In children living with HIV-1 (CLHIV), repeated dosing showed dose-dependent decreases in viral load (139). Despite these encouraging results, PRO-542 suffered from poor pharmacokinetics, with a serum half-life of about 2-3.5 days (compared to 18-21 days

typically expected for antibodies). Additionally, dose-escalation studies revealed nonlinear clearance kinetics, suggesting target-mediated drug disposition (TMDD) (138–140).

The concept of CD4-Ig was significantly advanced by the development of eCD4-Ig in the Farzan lab. In their construct, a sulfated CCR5-mimetic peptide (mim2) was fused to the C-terminus of the CD4-Ig Fc domain, creating a molecule capable of engaging both CD4 and the co-receptor binding sites simultaneously on gp120 (103). This dual-targeting design greatly improved both breadth and potency, as eCD4-Ig neutralized 100% of a diverse panel of neutralization-resistant HIV-1, HIV-2 and SHIV isolates with single-digit nanomolar to picomolar IC_{50} values. When delivered via adeno-associated virus (AAV) in rhesus macaques, eCD4-Ig provided durable protection against repeated intravenous SHIV challenge for at least 34 weeks. *In vitro* escape studies further show eCD4-Ig constrained viral escape more effectively than either CD4-Ig alone or VRC01 because mutations that reduced susceptibility to eCD4-Ig imposed substantial fitness costs to the virus (136).

Despite these promising results, the clinical development of CD4-Ig and eCD4-Ig has been limited by several interconnected issues associated with truncated D1D2 wild-type (WT) CD4 constructs. First, removal of the D3 and D4 domains, which are not required for HIV-1 engagement, exposes residues on the C-terminus of D2 that affect the construct's thermostability (referred to herein as the apparent melting temperature, T_m), likely contributing to protein aggregation and faster *in vivo* clearance (141). Second, these constructs retain binding to endogenous MHC class II molecules on antigen-presenting cells, creating an unintended sink that can drive TMDD and reduce the amount of drug available to neutralize the virus. These factors likely contribute to the short serum half-life observed for PRO-542. Overcoming these limitations through rational protein design is the focus of Chapter 3, described briefly in the following section.

1.6 Targeting Invariant Domains of HIV-1

1.6.1 The CD4 binding site: engineering gCD4 for improved CD4-Ig therapeutics

A major challenge in improving CD4 for therapeutic use is the fact that the binding footprints for gp120 and MHC-II overlap extensively. Both interactions occur primarily on the same face of the CD4 D1 domain, spanning residues 23-64, and include key residues such as F43 and R59 which contribute importantly to both interfaces (*142*). This overlap makes it difficult to disrupt the engineered CD4's unwanted interaction with MHC-II without also weakening the desired interaction with HIV-1 Env. Thankfully, the number of residues that are exclusive to MHC-II binding are few, but not absent. gp120 interacts with CD4 as a monomer on one side, whereas MHC-II is a heterodimer whose α and β chains engage two faces on the CD4 D1 domain. As a result, we find residues S60 and D63 contact the MHC-II α chain but does not participate in gp120 binding. This creates a small but exploitable “wedge” that can be engineered to selectively ablate MHC-II recognition without compromising gp120 affinity.

Our group took advantage of this asymmetry using a combined computational and structure-guided design strategy. An initial analysis by Adrian Bahn using PyMOL and TRIAD protein software helped identify a hydrophobic patch on the C-terminal region of the CD4 D2 domain formed by residues L109, L151, V175, and L177. In full-length CD4 (with four immunoglobulin-like domains, D1-D4), these side chains are partially buried through contacts with domains D3. However, in CD4-Ig, which contain only the D1D2 domains, this region becomes exposed, creating a thermodynamically unfavorable hydrophobic patch. *In silico* energy calculations were used to identify substitutions to these residues predicted to stabilize this exposed patch. Experimental screening ultimately identified the combination of mutations L109K, L151K, V175I, and L177D (KKID), which formed a favorable network of salt bridges and increased thermostability, T_m , by 6°C. An additional stabilizing mutation, A55V, was previously reported to show improved packing in the D1 hydrophobic core and, when included, contributed an additional 8°C increase in thermostability (*143*). The combined effects of these mutations were additive and together contributed a 14°C improvement.

In parallel, residues S60 and D63 were mutated to ablate MHC-II recognition. Guided by analyses of CD4 interacting with MHC-II using high-affinity Q40Y/T45W variants, site-directed mutagenesis identified S60D and D63K as optimal substitutions for disrupting contacts with the MHC-II α chain. Surface plasmon resonance (SPR) measurements confirmed these mutations indeed reduced MHC-II binding to near-background levels comparable to the negative control while preserving gp120 binding in the single-digit nanomolar range.

Combining both sets of stabilizing and MHC-II-specific mutations produced gp120-specific CD4, gCD4, with the final set of six key mutations including A55V, S60D, D63K, L109K, L151K, and L177E (VDKKKE). The L177E substitution was ultimately chosen in place of L177D because the longer glutamate side chain could support more favorable salt bridge formation. Relative to WT CD4 D1D2, gCD4 showed a significantly improved T_m of 66.5°C, more than 20°C above WT, yielded nearly 50-fold more protein in mammalian expression, and retained full gp120 binding while exhibiting near-complete loss of MHC-II recognition.

1.6.2 The caldera region: a novel conserved epitope in open HIV-1 Env

In a closed, prefusion Env trimer, the most highly conserved protein surfaces are buried underneath the glycan shield and/or the interfaces between the gp120-gp41 protomers, inaccessible to antibodies. CD4, however, changes this landscape. When engaged with CD4 (or a CD4-mimetic compound), the trimer undergoes conformational changes and enters an “open” state where the gp120 subunits are rotated outward, thereby exposing a previously hidden surface between the displaced gp120 protomers and above the gp41 three-helix bundle. Our group has termed this highly conserved, glycan-free region the caldera.

Cryo-EM studies of BG505 SOSIP trimers bound to soluble CD4 and the CD4-mimetic peptide M48U1 have characterized the caldera and shown that it is remarkably conserved across HIV-1 clades (86, 88, 89). Sequence analyses spanning thousands of HIV-1 isolates further show that caldera residues are among the most conserved on the entire Env surface. Notably, this region is distinct from the epitopes recognized by classical CD4-induced

(CD4i) antibodies, such as 17b, E51, and 21c, which have limited neutralization capabilities (144–147). These antibodies target the co-receptor binding site (including the bridging sheet and base of V3) on individual gp120 monomers, whereas the caldera is an inter-protomer surface created only after trimer opening (148).

The conservation of the caldera region reflects its strong functional constraints. In the closed trimer, these surfaces help stabilize the prefusion state through inter-protomer interactions. During viral entry, they must undergo tightly coordinated rearrangements to expose and position the fusion machinery correctly. Therefore, mutations in this region would be expected to destabilize Env and/or interfere with these necessary conformational changes. This imposes substantial fitness costs, thereby making the caldera region an attractive target for a pan-HIV-1 therapeutic.

An effective therapeutic directed at this region could be made through construction of a bi-functional molecule containing CD4/CD4-mimetic that would first induce conformational changes in Env, thereby exposing the caldera, before a second caldera-binding element engages the newly revealed epitope. By stabilizing Env in its “open” state while blocking the caldera, this construct could theoretically (1) block viral entry by preventing co-receptor engagement, (2) inactivate the spike by binding the gp41 six-helix bundle and inhibiting membrane fusion, (3) initiate Fc-mediated effector functions, and (4) resist escape at both the CD4bs and caldera epitopes.

1.6.3 Nanobody (V_{HH}) discovery to target the HIV-1 Env caldera

IgG antibodies are the well-established standard for conventional therapeutic applications. They are approximately 150 kDa in molecular weight and are made of a complex heterodimeric architecture (two heavy and two light chains) (149). Due to their relative size and complexity, they may face steric challenges in accessing the caldera epitope as its located in a recessed region between the gp120 protomers. The antigen-binding fragments (Fabs) may be excluded due to their relative bulkiness unless they have particularly long complementarity determining region 3 (CDR3) loops. Therefore, as a proof-of-principle, we

decided to pursue single-domain antibodies derived from the antigen-binding fragments of camelid heavy-chain-only antibodies—variable heavy-chain (V_{HH}) domains, or “nanobodies” (150, 151). Due to their small size (about 15 kDa, compared to roughly 50 kDa for Fabs), nanobodies are well-suited to overcome this obstacle, especially since their single-domain V_{HH} architecture makes them half as wide as a Fab’s V_H - V_L structure (**Figure 3**). Despite their smaller size, nanobodies retain full antigen-binding activity through their three CDR loops. Taking this all into account, nanobodies can potentially project effectively into narrow clefts and reach cryptic epitopes that are normally inaccessible to larger antibodies.

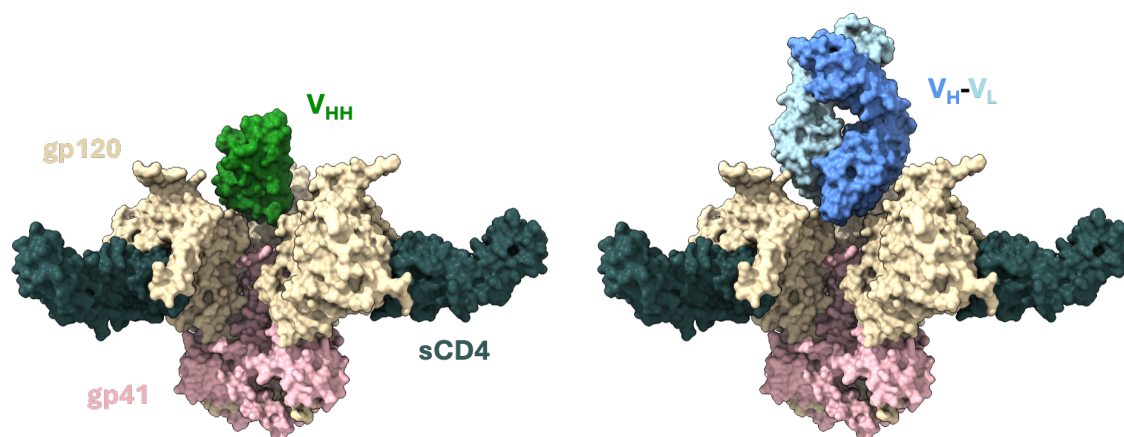


Figure 3. Nanobody domains (V_{HH} ; left) may be better suited than conventional antibody fragments (V_H - V_L Fabs; right) or single-chain Fvs (scFvs; single-chain V_H - V_L constructs, not shown) to access recessed epitopes on CD4-bound open HIV-1 Env. SOSIP-sCD4 complex from PDB: 5vn3. Representative J3 V_{HH} from PDB: 7lpn. Representative 17b Fab V_H - V_L from PDB: 1rz8.

Nanobodies can be produced efficiently and inexpensively in bacterial systems and be readily adapted for genetic fusion to other protein domains. In our case, a caldera-binding nanobody could be fused at its C-terminus to an IgG Fc and/or CD4 to create a bi-functional construct. Chapter 4 describes our efforts in the discovery and characterization of nanobodies that recognize HIV-1 Env exclusively in its “open” conformation. Briefly, nanobodies were identified using yeast display with a V_{HH} library against SOSIP trimers from multiple HIV-1 clades, including BG505 (clade A), B41 (clade B), and ZM197M (clade C) (152). The

screening strategy used “closed” DS-SOSIPs (which contain a 201C-433C disulfide bond that prevents CD4-induced opening) to deplete binders to the exterior of Env and followed with binder selection and enrichment using “open” SOSIP-CD4 complexes. With efforts ongoing, the hope is to discover a novel caldera-binding nanobody that can be used as part of a pan-HIV-1 therapeutic.

1.7 Scope of Work

This thesis brings together three projects for the common goal of developing broadly protective strategies against pandemic-level threats through molecular engineering. Chapter 2 examines computationally designed mosaic nanoparticle vaccines that generate broad immunity against SARS-CoV-2 and related sarbecoviruses with zoonotic potential. Chapters 3 and 4 then shift to the development of broad and potent therapeutics against HIV-1 that target two distinct conserved regions in Env. Together, these projects illustrate how targeting conserved viral features can enable the development of more durable countermeasures against highly mutable pathogens.

1.8 References

1. Centers for Disease Control (CDC), Pneumocystis pneumonia--Los Angeles. *MMWR Morb. Mortal. Wkly. Rep.* **30**, 250–252 (1981).
2. P. M. Sharp, B. H. Hahn, Origins of HIV and the AIDS Pandemic. *Cold Spring Harb. Perspect. Med.* **1**, a006841 (2011).
3. N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, P. Niu, F. Zhan, X. Ma, D. Wang, W. Xu, G. Wu, G. F. Gao, W. Tan, A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N. Engl. J. Med.* **382**, 727–733 (2020).
4. C. Wang, P. W. Horby, F. G. Hayden, G. F. Gao, A novel coronavirus outbreak of global health concern. *The Lancet* **395**, 470–473 (2020).
5. F. Gao, E. Bailes, D. L. Robertson, Y. Chen, C. M. Rodenburg, S. F. Michael, L. B. Cummins, L. O. Arthur, M. Peeters, G. M. Shaw, P. M. Sharp, B. H. Hahn, Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. *Nature* **397**, 436–441 (1999).
6. X. Jiang, R. Wang, Wildlife trade is likely the source of SARS-CoV-2. *Science* **377**, 925–926 (2022).
7. B. F. Keele, F. Van Heuverswyn, Y. Li, E. Bailes, J. Takehisa, M. L. Santiago, F. Bibollet-Ruche, Y. Chen, L. V. Wain, F. Liegeois, S. Loul, E. M. Ngole, Y. Bienvenue, E. Delaporte, J. F. Y. Brookfield, P. M. Sharp, G. M. Shaw, M. Peeters, B. H. Hahn, Chimpanzee Reservoirs of Pandemic and Nonpandemic HIV-1. *Science* **313**, 523–526 (2006).
8. Centers for Disease Control (CDC), Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men--New York City and California. *MMWR Morb. Mortal. Wkly. Rep.* **30**, 305–308 (1981).
9. U. R. UNAIDS Report, Global HIV & AIDS statistics — Fact sheet. <https://www.unaids.org/en/resources/fact-sheet>.
10. A. S. Fauci, HIV and AIDS: 20 years of science. *Nat. Med.* **9**, 839–843 (2003).
11. Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. M. Leung, E. H. Y. Lau, J. Y. Wong, X. Xing, N. Xiang, Y. Wu, C. Li, Q. Chen, D. Li, T. Liu, J. Zhao, M. Liu, W. Tu, C. Chen, L. Jin, R. Yang, Q. Wang, S. Zhou, R. Wang, H. Liu, Y. Luo, Y. Liu, G. Shao, H. Li, Z. Tao, Y. Yang, Z. Deng, B. Liu, Z. Ma, Y. Zhang, G. Shi, T. T. Y. Lam, J. T. Wu, G. F. Gao, B. J. Cowling, B. Yang, G. M. Leung, Z. Feng, Early

Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. *N. Engl. J. Med.* **382**, 1199–1207 (2020).

12. M. Worobey, J. I. Levy, L. Malpica Serrano, A. Crits-Christoph, J. E. Pekar, S. A. Goldstein, A. L. Rasmussen, M. U. G. Kraemer, C. Newman, M. P. G. Koopmans, M. A. Suchard, J. O. Wertheim, P. Lemey, D. L. Robertson, R. F. Garry, E. C. Holmes, A. Rambaut, K. G. Andersen, The Huanan Seafood Wholesale Market in Wuhan was the early epicenter of the COVID-19 pandemic. *Science* **377**, 951–959 (2022).
13. J. L. Havens, S. L. K. Pond, J. D. Zehr, J. E. Pekar, E. Parker, M. Worobey, K. G. Andersen, J. O. Wertheim, Dynamics of natural selection preceding human viral epidemics and pandemics. *Cell* **0** (2026).
14. A. Schuchat, Public Health Response to the Initiation and Spread of Pandemic COVID-19 in the United States, February 24–April 21, 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69** (2020).
15. H. Elser, M. V. Kiang, E. M. John, J. F. Simard, M. Bondy, L. M. Nelson, W. Chen, E. Linos, The Impact of the first COVID-19 shelter-in-place announcement on social distancing, difficulty in daily activities, and levels of concern in the San Francisco Bay Area: A cross-sectional social media survey. *PLOS ONE* **16**, e0244819 (2021).
16. E. Dong, H. Du, L. Gardner, An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* **20**, 533–534 (2020).
17. W. Msemburi, A. Karlinsky, V. Knutson, S. Aleshin-Guendel, S. Chatterji, J. Wakefield, The WHO estimates of excess mortality associated with the COVID-19 pandemic. *Nature* **613**, 130–137 (2023).
18. J. M. Cuevas, R. Geller, R. Garijo, J. López-Aldeguer, R. Sanjuán, Extremely High Mutation Rate of HIV-1 In Vivo. *PLOS Biol.* **13**, e1002251 (2015).
19. A. J. Greaney, T. N. Starr, P. Gilchuk, S. J. Zost, E. Binshtein, A. N. Loes, S. K. Hilton, J. Huddleston, R. Eguia, K. H. D. Crawford, A. S. Dingens, R. S. Nargi, R. E. Sutton, N. Suryadevara, P. W. Rothlauf, Z. Liu, S. P. J. Whelan, R. H. Carnahan, J. E. Crowe, J. D. Bloom, Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition. *Cell Host Microbe* **29**, 44-57.e9 (2021).
20. Y. Cao, J. Wang, F. Jian, T. Xiao, W. Song, A. Yisimayi, W. Huang, Q. Li, P. Wang, R. An, J. Wang, Y. Wang, X. Niu, S. Yang, H. Liang, H. Sun, T. Li, Y. Yu, Q. Cui, S. Liu, X. Yang, S. Du, Z. Zhang, X. Hao, F. Shao, R. Jin, X. Wang, J. Xiao, Y. Wang,

- X. S. Xie, Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies. *Nature* **602**, 657–663 (2022).
21. B. Choi, M. C. Choudhary, J. Regan, J. A. Sparks, R. F. Padera, X. Qiu, I. H. Solomon, H.-H. Kuo, J. Boucau, K. Bowman, U. D. Adhikari, M. L. Winkler, A. A. Mueller, T. Y.-T. Hsu, M. Desjardins, L. R. Baden, B. T. Chan, B. D. Walker, M. Lichterfeld, M. Brigl, D. S. Kwon, S. Kanjilal, E. T. Richardson, A. H. Jonsson, G. Alter, A. K. Barczak, W. P. Hanage, X. G. Yu, G. D. Gaiha, M. S. Seaman, M. Cernadas, J. Z. Li, Persistence and Evolution of SARS-CoV-2 in an Immunocompromised Host. *N. Engl. J. Med.* **383**, 2291–2293 (2020).
 22. R. Craigie, F. D. Bushman, HIV DNA Integration. *Cold Spring Harb. Perspect. Med.* **2**, a006890 (2012).
 23. J. D. Siliciano, R. F. Siliciano, A long-term latent reservoir for HIV-1: discovery and clinical implications. *J. Antimicrob. Chemother.* **54**, 6–9 (2004).
 24. D. Finzi, M. Hermankova, T. Pierson, L. M. Carruth, C. Buck, R. E. Chaisson, T. C. Quinn, K. Chadwick, J. Margolick, R. Brookmeyer, J. Gallant, M. Markowitz, D. D. Ho, D. D. Richman, R. F. Siliciano, Identification of a Reservoir for HIV-1 in Patients on Highly Active Antiretroviral Therapy. *Science* **278**, 1295–1300 (1997).
 25. T.-W. Chun, L. Stuyver, S. B. Mizell, L. A. Ehler, J. A. M. Mican, M. Baseler, A. L. Lloyd, M. A. Nowak, A. S. Fauci, Presence of an inducible HIV-1 latent reservoir during highly active antiretroviral therapy. *Proc. Natl. Acad. Sci.* **94**, 13193–13197 (1997).
 26. The Stages of HIV Infection | NIH. <https://hivinfo.nih.gov/understanding-hiv/fact-sheets/stages-hiv-infection>.
 27. X. Wei, J. M. Decker, S. Wang, H. Hui, J. C. Kappes, X. Wu, J. F. Salazar-Gonzalez, M. G. Salazar, J. M. Kilby, M. S. Saag, N. L. Komarova, M. A. Nowak, B. H. Hahn, P. D. Kwong, G. M. Shaw, Antibody neutralization and escape by HIV-1. *Nature* **422**, 307–312 (2003).
 28. D. D. Richman, T. Wrin, S. J. Little, C. J. Petropoulos, Rapid evolution of the neutralizing antibody response to HIV type 1 infection. *Proc. Natl. Acad. Sci.* **100**, 4144–4149 (2003).
 29. M. S. Cohen, Y. Q. Chen, M. McCauley, T. Gamble, M. C. Hosseini-pour, N. Kumarasamy, J. G. Hakim, J. Kumwenda, B. Grinsztejn, J. H. S. Pilotto, S. V. Godbole, S. Mehendale, S. Chariyalertsak, B. R. Santos, K. H. Mayer, I. F. Hoffman, S. H.

- Eshleman, E. Piwowar-Manning, L. Wang, J. Makhema, L. A. Mills, G. de Bruyn, I. Sanne, J. Eron, J. Gallant, D. Havlir, S. Swindells, H. Ribaudo, V. Elharrar, D. Burns, T. E. Taha, K. Nielsen-Saines, D. Celentano, M. Essex, T. R. Fleming, Prevention of HIV-1 Infection with Early Antiretroviral Therapy. *N. Engl. J. Med.* **365**, 493–505 (2011).
30. A. J. Rodger, V. Cambiano, T. Bruun, P. Vernazza, S. Collins, J. van Lunzen, G. M. Corbelli, V. Estrada, A. M. Geretti, A. Beloukas, D. Asboe, P. Viciano, F. Gutiérrez, B. Clotet, C. Pradier, J. Gerstoft, R. Weber, K. Westling, G. Wandeler, J. M. Prins, A. Rieger, M. Stoeckle, T. Kümmerle, T. Bini, A. Ammassari, R. Gilson, I. Krznaric, M. Ristola, R. Zangerle, P. Handberg, A. Antela, S. Allan, A. N. Phillips, J. Lundgren, for the PARTNER Study Group, Sexual Activity Without Condoms and Risk of HIV Transmission in Serodifferent Couples When the HIV-Positive Partner Is Using Suppressive Antiretroviral Therapy. *JAMA* **316**, 171–181 (2016).
 31. S. Sanche, Y. T. Lin, C. Xu, E. Romero-Severson, N. Hengartner, R. Ke, High Contagiousness and Rapid Spread of Severe Acute Respiratory Syndrome Coronavirus 2 - Volume 26, Number 7—July 2020 - Emerging Infectious Diseases journal - CDC. doi: 10.3201/eid2607.200282.
 32. X. He, E. H. Y. Lau, P. Wu, X. Deng, J. Wang, X. Hao, Y. C. Lau, J. Y. Wong, Y. Guan, X. Tan, X. Mo, Y. Chen, B. Liao, W. Chen, F. Hu, Q. Zhang, M. Zhong, Y. Wu, L. Zhao, F. Zhang, B. J. Cowling, F. Li, G. M. Leung, Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat. Med.* **26**, 672–675 (2020).
 33. Y. Liu, J. Rocklöv, The effective reproductive number of the Omicron variant of SARS-CoV-2 is several times relative to Delta. *J. Travel Med.* **29**, taac037 (2022).
 34. M. Hoffmann, H. Kleine-Weber, S. Schroeder, N. Krüger, T. Herrler, S. Erichsen, T. S. Schiergens, G. Herrler, N.-H. Wu, A. Nitsche, M. A. Müller, C. Drosten, S. Pöhlmann, SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **181**, 271-280.e8 (2020).
 35. E. J. Williamson, A. J. Walker, K. Bhaskaran, S. Bacon, C. Bates, C. E. Morton, H. J. Curtis, A. Mehrkar, D. Evans, P. Inglesby, J. Cockburn, H. I. McDonald, B. MacKenna, L. Tomlinson, I. J. Douglas, C. T. Rentsch, R. Mathur, A. Y. S. Wong, R. Grieve, D. Harrison, H. Forbes, A. Schultze, R. Croker, J. Parry, F. Hester, S. Harper, R. Perera, S. J. W. Evans, L. Smeeth, B. Goldacre, Factors associated with COVID-19-related death using OpenSAFELY. *Nature* **584**, 430–436 (2020).

36. E. Hartenian, D. Nandakumar, A. Lari, M. Ly, J. M. Tucker, B. A. Glaunsinger, The molecular virology of coronaviruses. *J. Biol. Chem.* **295**, 12910–12934 (2020).
37. 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).
38. M. Letko, A. Marzi, V. Munster, Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat. Microbiol.* **5**, 562–569 (2020).
39. A. C. Walls, Y.-J. Park, M. A. Tortorici, A. Wall, A. T. McGuire, D. Veasler, Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **181**, 281-292.e6 (2020).
40. Y. Guan, B. J. Zheng, Y. Q. He, X. L. Liu, Z. X. Zhuang, C. L. Cheung, S. W. Luo, P. H. Li, L. J. Zhang, Y. J. Guan, K. M. Butt, K. L. Wong, K. W. Chan, W. Lim, K. F. Shortridge, K. Y. Yuen, J. S. M. Peiris, L. L. M. Poon, Isolation and Characterization of Viruses Related to the SARS Coronavirus from Animals in Southern China. *Science* **302**, 276–278 (2003).
41. E. I. Azhar, S. A. El-Kafrawy, S. A. Farraj, A. M. Hassan, M. S. Al-Saeed, A. M. Hashem, T. A. Madani, Evidence for Camel-to-Human Transmission of MERS Coronavirus. *N. Engl. J. Med.* **370**, 2499–2505 (2014).
42. V. D. Menachery, B. L. Yount, A. C. Sims, K. Debbink, S. S. Agnihothram, L. E. Gralinski, R. L. Graham, T. Scobey, J. A. Plante, S. R. Royal, J. Swanstrom, T. P. Sheahan, R. J. Pickles, D. Corti, S. H. Randell, A. Lanzavecchia, W. A. Marasco, R. S. Baric, SARS-like WIV1-CoV poised for human emergence. *Proc. Natl. Acad. Sci.* **113**, 3048–3053 (2016).
43. A. A. Cohen, P. N. P. Gnanapragasam, Y. E. Lee, P. R. Hoffman, S. Ou, L. M. Kakutani, J. R. Keefe, H.-J. Wu, M. Howarth, A. P. West, C. O. Barnes, M. C. Nussenzweig, P. J. Bjorkman, Mosaic nanoparticles elicit cross-reactive immune responses to zoonotic coronaviruses in mice. *Science* **371**, 735–741 (2021).
44. A. A. Cohen, N. van Doremalen, A. J. Greaney, H. Andersen, A. Sharma, T. N. Starr, J. R. Keefe, C. Fan, J. E. Schulz, P. N. P. Gnanapragasam, L. M. Kakutani, A. P. West, G. Saturday, Y. E. Lee, H. Gao, C. A. Jette, M. G. Lewis, T. K. Tan, A. R. Townsend, J. D. Bloom, V. J. Munster, P. J. Bjorkman, Mosaic RBD nanoparticles protect against challenge by diverse sarbecoviruses in animal models. *Science* **377**, eabq0839 (2022).

45. M. F. Boni, P. Lemey, X. Jiang, T. T.-Y. Lam, B. W. Perry, T. A. Castoe, A. Rambaut, D. L. Robertson, Evolutionary origins of the SARS-CoV-2 sarbecovirus lineage responsible for the COVID-19 pandemic. *Nat. Microbiol.* **5**, 1408–1417 (2020).
46. D. Wrapp, N. Wang, K. S. Corbett, J. A. Goldsmith, C.-L. Hsieh, O. Abiona, B. S. Graham, J. S. McLellan, Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260–1263 (2020).
47. D. Bestle, M. R. Heindl, H. Limburg, T. V. L. Van, O. Pilgram, H. Moulton, D. A. Stein, K. Hards, M. Eickmann, O. Dolnik, C. Rohde, H.-D. Klenk, W. Garten, T. Steinmetzer, E. Böttcher-Friebertshäuser, TMPRSS2 and furin are both essential for proteolytic activation of SARS-CoV-2 in human airway cells. *Life Sci. Alliance* **3** (2020).
48. J. Lan, J. Ge, J. Yu, S. Shan, H. Zhou, S. Fan, Q. Zhang, X. Shi, Q. Wang, L. Zhang, X. Wang, Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **581**, 215–220 (2020).
49. R. Yan, Y. Zhang, Y. Li, L. Xia, Y. Guo, Q. Zhou, Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **367**, 1444–1448 (2020).
50. M. E. Abernathy, K.-M. A. Dam, S. R. Esswein, C. A. Jette, P. J. Bjorkman, How Antibodies Recognize Pathogenic Viruses: Structural Correlates of Antibody Neutralization of HIV-1, SARS-CoV-2, and Zika. *Viruses* **13** (2021).
51. C. O. Barnes, C. A. Jette, M. E. Abernathy, K.-M. A. Dam, S. R. Esswein, H. B. Gristick, A. G. Malyutin, N. G. Sharaf, K. E. Huey-Tubman, Y. E. Lee, D. F. Robbiani, M. C. Nussenzweig, A. P. West, P. J. Bjorkman, SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature* **588**, 682–687 (2020).
52. C. O. Barnes, A. P. West, K. E. Huey-Tubman, M. A. G. Hoffmann, N. G. Sharaf, P. R. Hoffman, N. Koranda, H. B. Gristick, C. Gaebler, F. Muecksch, J. C. C. Lorenzi, S. Finkin, T. Hägglöf, A. Hurley, K. G. Millard, Y. Weisblum, F. Schmidt, T. Hatziioannou, P. D. Bieniasz, M. Caskey, D. F. Robbiani, M. C. Nussenzweig, P. J. Bjorkman, Structures of Human Antibodies Bound to SARS-CoV-2 Spike Reveal Common Epitopes and Recurrent Features of Antibodies. *Cell* **182**, 828-842.e16 (2020).
53. D. Pinto, Y.-J. Park, M. Beltramello, A. C. Walls, M. A. Tortorici, S. Bianchi, S. Jaconi, K. Culap, F. Zatta, A. De Marco, A. Peter, B. Guarino, R. Spreafico, E. Cameroni, J. B. Case, R. E. Chen, C. Havenar-Daughton, G. Snell, A. Telenti, H. W. Virgin, A. Lanzavecchia, M. S. Diamond, K. Fink, D. Veessler, D. Corti, Cross-neutralization of

- SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature* **583**, 290–295 (2020).
54. C. A. Jette, A. A. Cohen, P. N. P. Gnanapragasam, F. Muecksch, Y. E. Lee, K. E. Huey-Tubman, F. Schmidt, T. Hatzioannou, P. D. Bieniasz, M. C. Nussenzweig, A. P. West, J. R. Keeffe, P. J. Bjorkman, C. O. Barnes, Broad cross-reactivity across sarbecoviruses exhibited by a subset of COVID-19 donor-derived neutralizing antibodies. *Cell Rep.* **36** (2021).
 55. C. Fan, A. A. Cohen, M. Park, A. F.-H. Hung, J. R. Keeffe, P. N. P. Gnanapragasam, Y. E. Lee, H. Gao, L. M. Kakutani, Z. Wu, H. Kleanthous, K. E. Malecek, J. C. Williams, P. J. Bjorkman, Neutralizing monoclonal antibodies elicited by mosaic RBD nanoparticles bind conserved sarbecovirus epitopes. *Immunity* **55**, 2419-2435.e10 (2022).
 56. M. S. Graham, C. H. Sudre, A. May, M. Antonelli, B. Murray, T. Varsavsky, K. Kläser, L. S. Canas, E. Molteni, M. Modat, D. A. Drew, L. H. Nguyen, L. Polidori, S. Selvachandran, C. Hu, J. Capdevila, C. Koshy, A. Ash, E. Wise, N. Moore, M. Mori, N. Cortes, J. Lynch, S. Kidd, D. J. Fairley, T. Curran, J. P. McKenna, H. Adams, C. Fraser, T. Golubchik, D. Bonsall, M. O. Hassan-Ibrahim, C. S. Malone, B. J. Cogger, M. Wantoch, N. Reynolds, B. Warne, J. Maksimovic, K. Spellman, K. McCluggage, M. John, R. Beer, S. Afifi, S. Morgan, A. Marchbank, A. Price, C. Kitchen, H. Gulliver, I. Merrick, J. Southgate, M. Guest, R. Munn, T. Workman, T. R. Connor, W. Fuller, C. Bresner, L. B. Snell, A. Patel, T. Charalampous, G. Nebbia, R. Batra, J. Edgeworth, S. C. Robson, A. H. Beckett, D. M. Aanensen, A. P. Underwood, C. A. Yeats, K. Abudahab, B. E. Taylor, M. Menegazzo, G. Clark, W. Smith, M. Khakh, V. M. Fleming, M. M. Lister, H. C. Howson-Wells, L. Berry, T. Boswell, A. Joseph, I. Willingham, C. Jones, C. Holmes, P. Bird, T. Helmer, K. Fallon, J. Tang, V. Raviprakash, S. Campbell, N. Sheriff, V. Blakey, L.-A. Williams, M. W. Loose, N. Holmes, C. Moore, M. Carlile, V. Wright, F. Sang, J. Debebe, F. Coll, A. W. Signell, G. Betancor, H. D. Wilson, S. Eldirdiri, A. Kenyon, T. Davis, O. G. Pybus, L. du Plessis, A. E. Zarebski, J. Raghwani, M. U. Kraemer, S. Francois, S. W. Attwood, T. I. Vasylyeva, M. E. Zamudio, B. Gutierrez, M. E. Torok, W. L. Hamilton, I. G. Goodfellow, G. Hall, A. S. Jahun, Y. Chaudhry, M. Hosmillo, M. L. Pinckert, I. Georgana, S. Moses, H. Lowe, L. Bedford, J. Moore, S. Stonehouse, C. L. Fisher, A. R. Awan, J. BoYes, J. Breuer, K. A. Harris, J. R. Brown, D. Shah, L. Atkinson, J. C. Lee, N. Storey, F. Flaviani, A. Alcolea-Medina, R. Williams, G. Vernet, M. R. Chapman, L. J. Levett, J. Heaney, W. Chatterton, M. Pusok, L. Xu-McCrae, D. L. Smith, M. Bashton, G. R. Young, A. Holmes, P. A. Randell, A. Cox, P. Madona, F. Bolt, J. Price, S. Mookerjee, M. Ragonnet-Cronin, F. F. Nascimento, D. Jorgensen, I.

Siveroni, R. Johnson, O. Boyd, L. Geidelberg, E. M. Volz, A. Rowan, G. P. Taylor, K. L. Smollett, N. J. Loman, J. Quick, C. McMurray, J. Stockton, S. Nicholls, W. Rowe, R. Poplawski, A. McNally, R. T. M. Nunez, J. Mason, T. I. Robinson, E. O'Toole, J. Watts, C. Breen, A. Cowell, G. Sluga, N. W. Machin, S. S. Y. Ahmad, R. P. George, F. Halstead, V. Sivaprakasam, W. Hogsden, C. J. Illingworth, C. Jackson, E. C. Thomson, J. G. Shepherd, P. Asamaphan, M. O. Niebel, K. K. Li, R. N. Shah, N. G. Jesudason, L. Tong, A. Broos, D. Mair, J. Nichols, S. N. Carmichael, K. Nomikou, E. Aranday-Cortes, N. Johnson, I. Starinskij, A. da S. Filipe, D. L. Robertson, R. J. Orton, J. Hughes, S. Vattipally, J. B. Singer, S. Nickbakhsh, A. D. Hale, L. R. Macfarlane-Smith, K. L. Harper, H. Carden, Y. Taha, B. A. Payne, S. Burton-Fanning, S. Waugh, J. Collins, G. Eltringham, S. Rushton, S. O'Brien, A. Bradley, A. Maclean, G. Mollett, R. Blacow, K. E. Templeton, M. P. McHugh, R. Dewar, E. Wastenge, S. Dervisevic, R. Stanley, E. J. Meader, L. Coupland, L. Smith, C. Graham, E. Barton, D. Padgett, G. Scott, E. Swindells, J. Greenaway, A. Nelson, C. M. McCann, W. C. Yew, M. Andersson, T. Peto, A. Justice, D. Eyre, D. Crook, T. J. Sloan, N. Duckworth, S. Walsh, A. J. Chauhan, S. Glaysher, K. Bicknell, S. Wyllie, S. Elliott, A. Lloyd, R. Impey, N. Levene, L. Monaghan, D. T. Bradley, T. Wyatt, E. Allara, C. Pearson, H. Osman, A. Bosworth, E. Robinson, P. Muir, I. B. Vipond, R. Hopes, H. M. Pymont, S. Hutchings, M. D. Curran, S. Parmar, A. Lackenby, T. Mbisa, S. Platt, S. Miah, D. Bibby, C. Manso, J. Hubb, M. Chand, G. Dabrera, M. Ramsay, D. Bradshaw, A. Thornton, R. Myers, U. Schaefer, N. Groves, E. Gallagher, D. Lee, D. Williams, N. Ellaby, I. Harrison, H. Hartman, N. Manesis, V. Patel, C. Bishop, V. Chalker, J. Ledesma, K. A. Twohig, M. T. G. Holden, S. Shaaban, A. Birchley, A. Adams, A. Davies, A. Gaskin, A. Plimmer, B. Gatica-Wilcox, C. McKerr, C. Moore, C. Williams, D. Heyburn, E. D. Lacy, E. Hilvers, F. Downing, G. Shankar, H. Jones, H. Asad, J. Coombes, J. Watkins, J. M. Evans, L. Fina, L. Gifford, L. Gilbert, L. Graham, M. Perry, M. Morgan, M. Bull, M. Cronin, N. Pacchiarini, N. Craine, R. Jones, R. Howe, S. Corden, S. Rey, S. Kumziene-SummerhaYes, S. Taylor, S. Cottrell, S. Jones, S. Edwards, J. O'Grady, A. J. Page, A. E. Mather, D. J. Baker, S. Rudder, A. Aydin, G. L. Kay, A. J. Trotter, N.-F. Alikhan, L. de O. Martins, T. Le-Viet, L. Meadows, A. Casey, L. Ratcliffe, D. A. Simpson, Z. Molnar, T. Thompson, E. Acheson, J. A. Masoli, B. A. Knight, S. Ellard, C. Auckland, C. R. Jones, T. W. Mahungu, D. Irish-Tavares, T. Haque, J. Hart, E. Witele, M. L. Fenton, A. Dadrah, A. Symmonds, T. Saluja, Y. Bourgeois, G. P. Scarlett, K. F. Loveson, S. Goudarzi, C. Fearn, K. Cook, H. Dent, H. Paul, D. G. Partridge, M. Raza, C. Evans, K. Johnson, S. Liggett, P. Baker, S. Bonner, S. Essex, R. A. Lyons, K. Saeed, A. I. K. Mahanama, B. Samaraweera, S. Silveira, E. Pelosi, E. Wilson-Davies, R. J. Williams, M. Kristiansen, S. Roy, C. A. Williams, M. Cotic, N. Bayzid, A. P. Westhorpe, J. A. Hartley, R. Jannoo, H. L. Lowe, A. Karamani, L. Ensell, J. A. Prieto, S. Jeremiah, D. Grammatopoulos, S. Pandey, L. Berry, K. Jones, A. Richter, A. Beggs, A. Best, B. Percival, J. Mirza, O. Megram, M. Mayhew, L. Crawford, F. Ashcroft, E.

- Moles-Garcia, N. Cumley, C. P. Smith, G. Bucca, A. R. Hesketh, B. Blane, S. T. Girgis, D. Leek, S. Sridhar, S. Forrest, C. Cormie, H. K. Gill, J. Dias, E. E. Higginson, M. Maes, J. Young, L. M. Kermack, R. K. Gupta, C. Ludden, S. J. Peacock, S. Palmer, C. M. Churcher, N. F. Hadjirin, A. M. Carabelli, E. Brooks, K. S. Smith, K. Galai, G. M. McManus, C. Ruis, R. K. Davidson, A. Rambaut, T. Williams, C. E. Balcazar, M. D. Gallagher, Á. O'Toole, S. Rooke, V. Hill, K. A. Williamson, T. D. Stanton, S. L. Michell, C. M. Bewshea, B. Temperton, M. L. Michelsen, J. Warwick-Dugdale, R. Manley, A. Farbos, J. W. Harrison, C. M. Sambles, D. J. Studholme, A. R. Jeffries, A. C. Darby, J. A. Hiscox, S. Paterson, M. Iturriza-Gomara, K. A. Jackson, A. O. Lucaci, E. E. Vamos, M. Hughes, L. Rainbow, R. Eccles, C. Nelson, M. Whitehead, L. Turtle, S. T. Haldenby, R. Gregory, M. Gemmell, C. Wierzbicki, H. J. Webster, T. I. de Silva, N. Smith, A. Angyal, B. B. Lindsey, D. C. Groves, L. R. Green, D. Wang, T. M. Freeman, M. D. Parker, A. J. Keeley, P. J. Parsons, R. M. Tucker, R. Brown, M. Wyles, M. Whiteley, P. Zhang, M. Gallis, S. F. Louka, C. Constantinidou, M. Unnikrishnan, S. Ott, J. K. J. Cheng, H. E. Bridgewater, L. R. Frost, G. Taylor-Joyce, R. Stark, L. Baxter, M. T. Alam, P. E. Brown, D. Aggarwal, A. C. Cerda, T. V. Merrill, R. E. Wilson, P. C. McClure, J. G. Chappell, T. Tsoleridis, J. Ball, D. Buck, J. A. Todd, A. Green, A. Trebes, G. MacIntyre-Cockett, M. de Cesare, A. Alderton, R. Amato, C. V. Ariani, M. A. Beale, C. Beaver, K. L. Bellis, E. Betteridge, J. Bonfield, J. Danesh, M. J. Dorman, E. Drury, B. W. Farr, L. Foulser, S. Goncalves, S. Goodwin, M. Gourtovaia, E. M. Harrison, D. K. Jackson, D. Jamrozy, I. Johnston, L. Kane, S. Kay, J.-P. Keatley, D. Kwiatkowski, C. F. Langford, M. Lawniczak, L. Letchford, R. Livett, S. Lo, I. Martincorena, S. McGuigan, R. Nelson, S. Palmer, N. R. Park, M. Patel, L. Prestwood, C. Puethe, M. A. Quail, S. Rajatileka, C. Scott, L. Shirley, J. Sillitoe, M. H. S. Chapman, S. A. Thurston, G. Tonkin-Hill, D. Weldon, D. Rajan, I. F. Bronner, L. Aigrain, N. M. Redshaw, S. V. Lensing, R. Davies, A. Whitwham, J. Liddle, K. Lewis, J. M. Tovar-Corona, S. Leonard, J. Durham, A. R. Bassett, S. McCarthy, R. J. Moll, K. James, K. Oliver, A. Makunin, J. Barrett, R. N. Gunson, A. Hammers, A. T. Chan, J. Wolf, T. D. Spector, C. J. Steves, S. Ourselin, Changes in symptomatology, reinfection, and transmissibility associated with the SARS-CoV-2 variant B.1.1.7: an ecological study. *Lancet Public Health* **6**, e335–e345 (2021).
57. H. Tegally, E. Wilkinson, M. Giovanetti, A. Iranzadeh, V. Fonseca, J. Giandhari, D. Doolabh, S. Pillay, E. J. San, N. Msomi, K. Mlisana, A. von Gottberg, S. Walaza, M. Allam, A. Ismail, T. Mohale, A. J. Glass, S. Engelbrecht, G. Van Zyl, W. Preiser, F. Petruccione, A. Sigal, D. Hardie, G. Marais, N. Hsiao, S. Korsman, M.-A. Davies, L. Tyers, I. Mudau, D. York, C. Maslo, D. Goedhals, S. Abrahams, O. Laguda-Akingba, A. Alisoltani-Dehkordi, A. Godzik, C. K. Wibmer, B. T. Sewell, J. Lourenço, L. C. J. Alcantara, S. L. Kosakovsky Pond, S. Weaver, D. Martin, R. J. Lessells, J. N. Bhiman,

- C. Williamson, T. de Oliveira, Detection of a SARS-CoV-2 variant of concern in South Africa. *Nature* **592**, 438–443 (2021).
58. D. Planas, D. Veyer, A. Baidaliuk, I. Staropoli, F. Guivel-Benhassine, M. M. Rajah, C. Planchais, F. Porrot, N. Robillard, J. Puech, M. Prot, F. Gallais, P. Gantner, A. Velay, J. Le Guen, N. Kassis-Chikhani, D. Edriss, L. Belec, A. Seve, L. Courtellemont, H. Péré, L. Hocqueloux, S. Fafi-Kremer, T. Prazuck, H. Mouquet, T. Bruel, E. Simon-Lorière, F. A. Rey, O. Schwartz, Reduced sensitivity of SARS-CoV-2 variant Delta to antibody neutralization. *Nature* **596**, 276–280 (2021).
 59. S. Cele, L. Jackson, D. S. Khoury, K. Khan, T. Moyo-Gwete, H. Tegally, J. E. San, D. Cromer, C. Scheepers, D. G. Amoako, F. Karim, M. Bernstein, G. Lustig, D. Archary, M. Smith, Y. Ganga, Z. Jule, K. Reedoy, S.-H. Hwa, J. Giandhari, J. M. Blackburn, B. I. Gosnell, S. S. Abdool Karim, W. Hanekom, A. von Gottberg, J. N. Bhiman, R. J. Lessells, M.-Y. S. Moosa, M. P. Davenport, T. de Oliveira, P. L. Moore, A. Sigal, Omicron extensively but incompletely escapes Pfizer BNT162b2 neutralization. *Nature* **602**, 654–656 (2022).
 60. L. Liu, S. Iketani, Y. Guo, J. F.-W. Chan, M. Wang, L. Liu, Y. Luo, H. Chu, Y. Huang, M. S. Nair, J. Yu, K. K.-H. Chik, T. T.-T. Yuen, C. Yoon, K. K.-W. To, H. Chen, M. T. Yin, M. E. Sobieszczyk, Y. Huang, H. H. Wang, Z. Sheng, K.-Y. Yuen, D. D. Ho, Striking antibody evasion manifested by the Omicron variant of SARS-CoV-2. *Nature* **602**, 676–681 (2022).
 61. T. N. Starr, A. J. Greaney, A. Addetia, W. W. Hannon, M. C. Choudhary, A. S. Dingens, J. Z. Li, J. D. Bloom, Prospective mapping of viral mutations that escape antibodies used to treat COVID-19. *Science* **371**, 850–854 (2021).
 62. N. P. Hachmann, J. Miller, A. Y. Collier, J. D. Ventura, J. Yu, M. Rowe, E. A. Bondzie, O. Powers, N. Surve, K. Hall, D. H. Barouch, Neutralization Escape by SARS-CoV-2 Omicron Subvariants BA.2.12.1, BA.4, and BA.5. *N. Engl. J. Med.* **387**, 86–88 (2022).
 63. F. P. Polack, S. J. Thomas, N. Kitchin, J. Absalon, A. Gurtman, S. Lockhart, J. L. Perez, G. P. Marc, E. D. Moreira, C. Zerbini, R. Bailey, K. A. Swanson, S. Roychoudhury, K. Koury, P. Li, W. V. Kalina, D. Cooper, R. W. Frenck, L. L. Hammitt, Ö. Türeci, H. Nell, A. Schaefer, S. Ünal, D. B. Tresnan, S. Mather, P. R. Dormitzer, U. Şahin, K. U. Jansen, W. C. Gruber, Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *N. Engl. J. Med.* **383**, 2603–2615 (2020).
 64. L. R. Baden, H. M. E. Sahly, B. Essink, K. Kotloff, S. Frey, R. Novak, D. Diemert, S. A. Spector, N. Rouphael, C. B. Creech, J. McGettigan, S. Khetan, N. Segall, J. Solis,

- A. Brosz, C. Fierro, H. Schwartz, K. Neuzil, L. Corey, P. Gilbert, H. Janes, D. Follmann, M. Marovich, J. Mascola, L. Polakowski, J. Ledgerwood, B. S. Graham, H. Bennett, R. Pajon, C. Knightly, B. Leav, W. Deng, H. Zhou, S. Han, M. Ivarsson, J. Miller, T. Zaks, Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. *N. Engl. J. Med.* **384**, 403–416 (2021).
65. M. W. Tenforde, W. H. Self, K. Adams, M. Gaglani, A. A. Ginde, T. McNeal, S. Ghamande, D. J. Douin, H. K. Talbot, J. D. Casey, N. M. Mohr, A. Zepeski, N. I. Shapiro, K. W. Gibbs, D. C. Files, D. N. Hager, A. Shehu, M. E. Prekker, H. L. Erickson, M. C. Exline, M. N. Gong, A. Mohamed, D. J. Henning, J. S. Steingrub, I. D. Peltan, S. M. Brown, E. T. Martin, A. S. Monto, A. Khan, C. L. Hough, L. W. Busse, C. C. ten Lohuis, A. Duggal, J. G. Wilson, A. J. Gordon, N. Qadir, S. Y. Chang, C. Mallow, C. Rivas, H. M. Babcock, J. H. Kwon, N. Halasa, J. D. Chappell, A. S. Luring, C. G. Grijalva, T. W. Rice, I. D. Jones, W. B. Stubblefield, A. Baughman, K. N. Womack, J. P. Rhoads, C. J. Lindsell, K. W. Hart, Y. Zhu, S. M. Olson, M. Kobayashi, J. R. Verani, M. M. Patel, Influenza and Other Viruses in the Acutely Ill (IVY) Network, Association Between mRNA Vaccination and COVID-19 Hospitalization and Disease Severity. *JAMA* **326**, 2043–2054 (2021).
 66. P. Tang, M. R. Hasan, H. Chemaitelly, H. M. Yassine, F. M. Benslimane, H. A. Al Khatib, S. AlMukdad, P. Coyle, H. H. Ayoub, Z. Al Kanaani, E. Al Kuwari, A. Jeremijenko, A. H. Kaleeckal, A. N. Latif, R. M. Shaik, H. F. Abdul Rahim, G. K. Nasrallah, M. G. Al Kuwari, H. E. Al Romaihi, A. A. Butt, M. H. Al-Thani, A. Al Khal, R. Bertollini, L. J. Abu-Raddad, BNT162b2 and mRNA-1273 COVID-19 vaccine effectiveness against the SARS-CoV-2 Delta variant in Qatar. *Nat. Med.* **27**, 2136–2143 (2021).
 67. H. G. Rosenblum, Interim Recommendations from the Advisory Committee on Immunization Practices for the Use of Bivalent Booster Doses of COVID-19 Vaccines — United States, October 2022. *MMWR Morb. Mortal. Wkly. Rep.* **71** (2022).
 68. S. F. Andrews, Y. Huang, K. Kaur, L. I. Popova, I. Y. Ho, N. T. Pauli, C. J. H. Dunand, W. M. Taylor, S. Lim, M. Huang, X. Qu, J.-H. Lee, M. Salgado-Ferrer, F. Krammer, P. Palese, J. Wrammert, R. Ahmed, P. C. Wilson, Immune history profoundly affects broadly protective B cell responses to influenza. *Sci. Transl. Med.* **7**, 316ra192-316ra192 (2015).
 69. L. Mesin, A. Schiepers, J. Ersching, A. Barbulescu, C. B. Cavazzoni, A. Angelini, T. Okada, T. Kurosaki, G. D. Vitoria, Restricted Clonality and Limited Germinal Center Reentry Characterize Memory B Cell Reactivation by Boosting. *Cell* **180**, 92-106.e11 (2020).

70. A. E. Powell, K. Zhang, M. Sanyal, S. Tang, P. A. Weidenbacher, S. Li, T. D. Pham, J. E. Pak, W. Chiu, P. S. Kim, A Single Immunization with Spike-Functionalized Ferritin Vaccines Elicits Neutralizing Antibody Responses against SARS-CoV-2 in Mice. *ACS Cent. Sci.* **7**, 183–199 (2021).
71. P. A.-B. Weidenbacher, M. Sanyal, N. Friedland, S. Tang, P. S. Arunachalam, M. Hu, O. S. Kumru, M. K. Morris, J. Fontenot, L. Shirreff, J. Do, Y.-C. Cheng, G. Vasudevan, M. B. Feinberg, F. J. Villinger, C. Hanson, S. B. Joshi, D. B. Volkin, B. Pulendran, P. S. Kim, A ferritin-based COVID-19 nanoparticle vaccine that elicits robust, durable, broad-spectrum neutralizing antisera in non-human primates. *Nat. Commun.* **14**, 2149 (2023).
72. T. U. J. Bruun, A.-M. C. Andersson, S. J. Draper, M. Howarth, Engineering a Rugged Nanoscaffold To Enhance Plug-and-Display Vaccination. *ACS Nano* **12**, 8855–8866 (2018).
73. A. A. Cohen, J. R. Keeffe, A. Schiepers, S. E. Dross, A. J. Greaney, A. V. Rorick, H. Gao, P. N. P. Gnanapragasam, C. Fan, A. P. West, A. I. Ramsingh, J. H. Erasmus, J. D. Pata, H. Muramatsu, N. Pardi, P. J. C. Lin, S. Baxter, R. Cruz, M. Quintanar-Audelo, E. Robb, C. Serrano-Amatriain, L. Magneschi, I. G. Fotheringham, D. H. Fuller, G. D. Victora, P. J. Bjorkman, Mosaic sarbecovirus nanoparticles elicit cross-reactive responses in pre-vaccinated animals. *Cell* **187**, 5554-5571.e19 (2024).
74. E. Wang, A. A. Cohen, L. F. Caldera, J. R. Keeffe, A. V. Rorick, Y. M. Adia, P. N. P. Gnanapragasam, P. J. Bjorkman, A. K. Chakraborty, Designed mosaic nanoparticles enhance cross-reactive immune responses in mice. *Cell* **188**, 1036-1050.e11 (2025).
75. T. N. Starr, A. J. Greaney, S. K. Hilton, D. Ellis, K. H. D. Crawford, A. S. Dingens, M. J. Navarro, J. E. Bowen, M. A. Tortorici, A. C. Walls, N. P. King, D. Veelsler, J. D. Bloom, Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell* **182**, 1295-1310.e20 (2020).
76. O. Conchillo-Solé, N. S. de Groot, F. X. Avilés, J. Vendrell, X. Daura, S. Ventura, AGGRESKAN: a server for the prediction and evaluation of “hot spots” of aggregation in polypeptides. *BMC Bioinformatics* **8**, 65 (2007).
77. F. Barré-Sinoussi, A. L. Ross, J.-F. Delfraissy, Past, present and future: 30 years of HIV research. *Nat. Rev. Microbiol.* **11**, 877–883 (2013).
78. P. D. Kwong, R. Wyatt, J. Robinson, R. W. Sweet, J. Sodroski, W. A. Hendrickson, Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* **393**, 648–659 (1998).

79. P. Prabakaran, A. S. Dimitrov, T. R. Fouts, D. S. Dimitrov, “Structure and Function of the HIV Envelope Glycoprotein as Entry Mediator, Vaccine Immunogen, and Target for Inhibitors” in *Advances in Pharmacology* (Academic Press, 2007; <https://www.sciencedirect.com/science/article/pii/S1054358907550027>) vol. 55, pp. 33–97.
80. J. S. Klein, P. J. Bjorkman, Few and Far Between: How HIV May Be Evading Antibody Avidity. *PLOS Pathog.* **6**, e1000908 (2010).
81. P. Zhu, J. Liu, J. Bess, E. Chertova, J. D. Lifson, H. Grisé, G. A. Ofek, K. A. Taylor, K. H. Roux, Distribution and three-dimensional structure of AIDS virus envelope spikes. *Nature* **441**, 847–852 (2006).
82. J.-P. Julien, A. Cupo, D. Sok, R. L. Stanfield, D. Lyumkis, M. C. Deller, P.-J. Klasse, D. R. Burton, R. W. Sanders, J. P. Moore, A. B. Ward, I. A. Wilson, Crystal Structure of a Soluble Cleaved HIV-1 Envelope Trimer. *Science* **342**, 1477–1483 (2013).
83. K. J. Doores, The HIV glycan shield as a target for broadly neutralizing antibodies. *FEBS J.* **282**, 4679–4691 (2015).
84. G. B. E. Stewart-Jones, C. Soto, T. Lemmin, G.-Y. Chuang, A. Druz, R. Kong, P. V. Thomas, K. Wagh, T. Zhou, A.-J. Behrens, T. Bylund, C. W. Choi, J. R. Davison, I. S. Georgiev, M. G. Joyce, Y. D. Kwon, M. Pancera, J. Taft, Y. Yang, B. Zhang, S. S. Shivatare, V. S. Shivatare, C.-C. D. Lee, C.-Y. Wu, C. A. Bewley, D. R. Burton, W. C. Koff, M. Connors, M. Crispin, U. Baxa, B. T. Korber, C.-H. Wong, J. R. Mascola, P. D. Kwong, Trimeric HIV-1-Env Structures Define Glycan Shields from Clades A, B, and G. *Cell* **165**, 813–826 (2016).
85. M. Crispin, A. B. Ward, I. A. Wilson, Structure and Immune Recognition of the HIV Glycan Shield. *Annu. Rev. Biophys.* **47**, 499–523 (2018).
86. H. Wang, A. A. Cohen, R. P. Galimidi, H. B. Gristick, G. J. Jensen, P. J. Bjorkman, Cryo-EM structure of a CD4-bound open HIV-1 envelope trimer reveals structural rearrangements of the gp120 V1V2 loop. *Proc. Natl. Acad. Sci.* **113**, E7151–E7158 (2016).
87. G. Ozorowski, J. Pallesen, N. de Val, D. Lyumkis, C. A. Cottrell, J. L. Torres, J. Copps, R. L. Stanfield, A. Cupo, P. Pugach, J. P. Moore, I. A. Wilson, A. B. Ward, Open and closed structures reveal allostery and pliability in the HIV-1 envelope spike. *Nature* **547**, 360–363 (2017).

88. C. A. Jette, C. O. Barnes, S. M. Kirk, B. Melillo, A. B. Smith, P. J. Bjorkman, Cryo-EM structures of HIV-1 trimer bound to CD4-mimetics BNM-III-170 and M48U1 adopt a CD4-bound open conformation. *Nat. Commun.* **12**, 1950 (2021).
89. Z. Yang, H. Wang, A. Z. Liu, H. B. Gristick, P. J. Bjorkman, Asymmetric opening of HIV-1 Env bound to CD4 and a coreceptor-mimicking antibody. *Nat. Struct. Mol. Biol.* **26**, 1167–1175 (2019).
90. B. M. Stadtmueller, M. D. Bridges, K.-M. Dam, M. T. Lerch, K. E. Huey-Tubman, W. L. Hubbell, P. J. Bjorkman, DEER Spectroscopy Measurements Reveal Multiple Conformations of HIV-1 SOSIP Envelopes that Show Similarities with Envelopes on Native Virions. *Immunity* **49**, 235-246.e4 (2018).
91. C. A. Janeway, The co-receptor function of CD4. *Semin. Immunol.* **3**, 153–160 (1991).
92. R. Zamoyska, CD4 and CD8: modulators of T-cell receptor recognition of antigen and of immune responses? *Curr. Opin. Immunol.* **10**, 82–87 (1998).
93. A. G. Dalgleish, P. C. L. Beverley, P. R. Clapham, D. H. Crawford, M. F. Greaves, R. A. Weiss, The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature* **312**, 763–767 (1984).
94. D. Klatzmann, E. Champagne, S. Chamaret, J. Gruest, D. Guetard, T. Hercend, J.-C. Gluckman, L. Montagnier, T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. *Nature* **312**, 767–768 (1984).
95. H. Choe, M. Farzan, Y. Sun, N. Sullivan, B. Rollins, P. D. Ponath, L. Wu, C. R. Mackay, G. LaRosa, W. Newman, N. Gerard, C. Gerard, J. Sodroski, The β -Chemokine Receptors CCR3 and CCR5 Facilitate Infection by Primary HIV-1 Isolates. *Cell* **85**, 1135–1148 (1996).
96. T. Dragic, V. Litwin, G. P. Allaway, S. R. Martin, Y. Huang, K. A. Nagashima, C. Cayanan, P. J. Maddon, R. A. Koup, J. P. Moore, W. A. Paxton, HIV-1 entry into CD4⁺ cells is mediated by the chemokine receptor CC-CKR-5. *Nature* **381**, 667–673 (1996).
97. H. Deng, R. Liu, W. Ellmeier, S. Choe, D. Unutmaz, M. Burkhart, P. D. Marzio, S. Marmon, R. E. Sutton, C. M. Hill, C. B. Davis, S. C. Peiper, T. J. Schall, D. R. Littman, N. R. Landau, Identification of a major co-receptor for primary isolates of HIV-1. *Nature* **381**, 661–666 (1996).
98. D. C. Chan, D. Fass, J. M. Berger, P. S. Kim, Core Structure of gp41 from the HIV Envelope Glycoprotein. *Cell* **89**, 263–273 (1997).

99. D. C. Chan, P. S. Kim, HIV Entry and Its Inhibition. *Cell* **93**, 681–684 (1998).
100. P. Dorr, M. Westby, S. Dobbs, P. Griffin, B. Irvine, M. Macartney, J. Mori, G. Rickett, C. Smith-Burchnell, C. Napier, R. Webster, D. Armour, D. Price, B. Stammen, A. Wood, M. Perros, Maraviroc (UK-427,857), a Potent, Orally Bioavailable, and Selective Small-Molecule Inhibitor of Chemokine Receptor CCR5 with Broad-Spectrum Anti-Human Immunodeficiency Virus Type 1 Activity. *Antimicrob. Agents Chemother.* **49**, 4721–4732 (2005).
101. J. P. Lalezari, K. Henry, M. O’Hearn, J. S. G. Montaner, P. J. Piliero, B. Trottier, S. Walmsley, C. Cohen, D. R. Kuritzkes, J. J. Eron, J. Chung, R. DeMasi, L. Donatucci, C. Drobnes, J. Delehanty, M. Salgo, Enfuvirtide, an HIV-1 Fusion Inhibitor, for Drug-Resistant HIV Infection in North and South America. *N. Engl. J. Med.* **348**, 2175–2185 (2003).
102. T. Zhou, I. Georgiev, X. Wu, Z.-Y. Yang, K. Dai, A. Finzi, Y. Do Kwon, J. F. Scheid, W. Shi, L. Xu, Y. Yang, J. Zhu, M. C. Nussenzweig, J. Sodroski, L. Shapiro, G. J. Nabel, J. R. Mascola, P. D. Kwong, Structural Basis for Broad and Potent Neutralization of HIV-1 by Antibody VRC01. *Science* **329**, 811–817 (2010).
103. M. R. Gardner, L. M. Kattenhorn, H. R. Kondur, M. von Schaewen, T. Dorfman, J. J. Chiang, K. G. Haworth, J. M. Decker, M. D. Alpert, C. C. Bailey, E. S. Neale, C. H. Fellinger, V. R. Joshi, S. P. Fuchs, J. M. Martinez-Navio, B. D. Quinlan, A. Y. Yao, H. Mouquet, J. Gorman, B. Zhang, P. Poignard, M. C. Nussenzweig, D. R. Burton, P. D. Kwong, M. Piatak, J. D. Lifson, G. Gao, R. C. Desrosiers, D. T. Evans, B. H. Hahn, A. Ploss, P. M. Cannon, M. S. Seaman, M. Farzan, AAV-expressed eCD4-Ig provides durable protection from multiple SHIV challenges. *Nature* **519**, 87–91 (2015).
104. B. Korber, B. Gaschen, K. Yusim, R. Thakallapally, C. Kesmir, V. Detours, Evolutionary and immunological implications of contemporary HIV-1 variation. *Br. Med. Bull.* **58**, 19–42 (2001).
105. J. Hemelaar, E. Gouws, P. D. Ghys, S. Osmanov, Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *AIDS* **20**, W13 (2006).
106. J. Hemelaar, R. Elangovan, J. Yun, L. Dickson-Tetteh, I. Fleminger, S. Kirtley, B. Williams, E. Gouws-Williams, P. D. Ghys, A. G. Abimiku, S. Agwale, C. Archibald, B. Avidor, M. G. Barbás, F. Barre-Sinoussi, B. Barugahare, E. H. Belabbes, S. Bertagnolio, D. Bix, A. F. Bobkov, J. Brandful, H. Bredell, C. A. Brennan, J. Brooks, M. Bruckova, L. Buonaguro, F. Buonaguro, S. Buttò, A. Buve, M. Campbell, J. Carr, A. Carrera, M. G. Carrillo, C. Celum, B. Chaplin, M. Charles, D. Chatzidimitriou, Z.

- Chen, K. Chijiwa, D. Cooper, P. Cunningham, A. Dagnra, C. F. de Gascun, J. D. Amo, E. Delgado, U. Dietrich, D. Dwyer, D. Ellenberger, B. Ensoli, M. Essex, F. Gao, H. Fleury, P. N. Fonjungo, V. Foulongne, D. A. Gadkari, F. García, R. Garsia, G. M. Gershy-Damet, J. R. Glynn, R. Goodall, Z. Grossman, M. Lindenmeyer-Guimarães, B. Hahn, R. L. Hamers, O. Hamouda, R. Handema, X. He, J. Herbeck, D. D. Ho, A. Holguin, M. Hosseinipour, G. Hunt, M. Ito, M. A. B. H. Kacem, E. Kahle, P. K. Kaleebu, M. Kalish, A. Kamarulzaman, C. Kang, P. Kanki, E. Karamov, J.-C. Karasi, K. Kayitenkore, T. Kelleher, D. Kitayaporn, L. G. Kostrikis, C. Kucherer, C. Lara, T. Leitner, K. Liitsola, J. Lingappa, M. Linka, I. L. de Rivera, V. Lukashov, S. Maayan, L. Mayr, F. McCutchan, N. Meda, E. Menu, F. Mhalu, D. Mloka, J. L. Mokili, B. Montes, O. Mor, M. Morgado, F. Mosha, A. Moussi, J. Mullins, R. Najera, M. Nasr, N. Ndemi, J. R. Neilson, V. R. Nerurkar, F. Neuhaus, C. Nolte, V. Novitsky, P. Nyambi, M. Ofner, F. J. Paladin, A. Papa, J. Pape, N. Parkin, C. Parry, M. Peeters, A. Pelletier, L. Pérez-Álvarez, D. Pillay, A. Pinto, T. D. Quang, C. Rademeyer, F. Raikanikoda, M. A. Rayfield, J.-M. Reynes, T. R. de Wit, K. E. Robbins, M. Rolland, C. Rousseau, J. Salazar-Gonzales, H. Salem, M. Salminen, H. Salomon, P. Sandstrom, M. L. Santiago, A. D. Sarr, B. Schroeder, M. Segondy, P. Selhorst, S. Sempala, J. Servais, A. Shaik, Y. Shao, A. Slim, M. A. Soares, E. Songok, D. Stewart, J. Stokes, S. Subbarao, R. Sutthent, J. Takehisa, A. Tanuri, K. K. Tee, K. Thapa, M. Thomson, T. Tran, W. Urassa, H. Ushijima, P. van de Perre, G. van der Groen, K. van Laethem, J. van Oosterhout, A. van Sighem, E. van Wijngaerden, A.-M. Vandamme, J. Vercauteren, N. Vidal, L. Wallace, C. Williamson, D. Wolday, J. Xu, C. Yang, L. Zhang, R. Zhang, Global and regional molecular epidemiology of HIV-1, 1990–2015: a systematic review, global survey, and trend analysis. *Lancet Infect. Dis.* **19**, 143–155 (2019).
107. B. Gaschen, J. Taylor, K. Yusim, B. Foley, F. Gao, D. Lang, V. Novitsky, B. Haynes, B. H. Hahn, T. Bhattacharya, B. Korber, Diversity Considerations in HIV-1 Vaccine Selection. *Science* **296**, 2354–2360 (2002).
108. A. S. Perelson, A. U. Neumann, M. Markowitz, J. M. Leonard, D. D. Ho, HIV-1 Dynamics in Vivo: Virion Clearance Rate, Infected Cell Life-Span, and Viral Generation Time. *Science* **271**, 1582–1586 (1996).
109. A. Rios, Fundamental challenges to the development of a preventive HIV vaccine. *Curr. Opin. Virol.* **29**, 26–32 (2018).
110. J. Esparza, A brief history of the global effort to develop a preventive HIV vaccine. *Vaccine* **31**, 3502–3518 (2013).

111. J. Kim, S. Vasan, J. H. Kim, J. A. Ake, Current approaches to HIV vaccine development: a narrative review. *J. Int. AIDS Soc.* **24**, e25793 (2021).
112. R. W. Sanders, R. Derking, A. Cupo, J.-P. Julien, A. Yasmeen, N. de Val, H. J. Kim, C. Blattner, A. T. de la Peña, J. Korzun, M. Golabek, K. de los Reyes, T. J. Ketas, M. J. van Gils, C. R. King, I. A. Wilson, A. B. Ward, P. J. Klasse, J. P. Moore, A Next-Generation Cleaved, Soluble HIV-1 Env Trimer, BG505 SOSIP.664 gp140, Expresses Multiple Epitopes for Broadly Neutralizing but Not Non-Neutralizing Antibodies. *PLOS Pathog.* **9**, e1003618 (2013).
113. P. Pugach, G. Ozorowski, A. Cupo, R. Ringe, A. Yasmeen, N. de Val, R. Derking, H. J. Kim, J. Korzun, M. Golabek, K. de los Reyes, T. J. Ketas, J.-P. Julien, D. R. Burton, I. A. Wilson, R. W. Sanders, P. J. Klasse, A. B. Ward, J. P. Moore, A Native-Like SOSIP.664 Trimer Based on an HIV-1 Subtype B env Gene. *J. Virol.* **89**, 3380–3395 (2015).
114. J.-P. Julien, J. H. Lee, G. Ozorowski, Y. Hua, A. Torrents de la Peña, S. W. de Taeye, T. Nieuwsma, A. Cupo, A. Yasmeen, M. Golabek, P. Pugach, P. J. Klasse, J. P. Moore, R. W. Sanders, A. B. Ward, I. A. Wilson, Design and structure of two HIV-1 clade C SOSIP.664 trimers that increase the arsenal of native-like Env immunogens. *Proc. Natl. Acad. Sci.* **112**, 11947–11952 (2015).
115. S. Rerks-Ngarm, P. Pitisuttithum, S. Nitayaphan, J. Kaewkungwal, J. Chiu, R. Paris, N. Prensri, C. Namwat, M. de Souza, E. Adams, M. Benenson, S. Gurunathan, J. Tartaglia, J. G. McNeil, D. P. Francis, D. Stablein, D. L. Birx, S. Chunsuttiwat, C. Khamboonruang, P. Thongcharoen, M. L. Robb, N. L. Michael, P. Kunasol, J. H. Kim, Vaccination with ALVAC and AIDSVAX to Prevent HIV-1 Infection in Thailand. *N. Engl. J. Med.* **361**, 2209–2220 (2009).
116. G. E. Gray, L.-G. Bekker, F. Laher, M. Malahleha, M. Allen, Z. Moodie, N. Grunenberg, Y. Huang, D. Grove, B. Prigmore, J. J. Kee, D. Benkeser, J. Hural, C. Innes, E. Lazarus, G. Meintjes, N. Naicker, D. Kalonji, M. Nchabeleng, M. Sebe, N. Singh, P. Kotze, S. Kassim, T. Dubula, V. Naicker, W. Brumskine, C. N. Ncayiya, A. M. Ward, N. Garrett, G. Kistnasami, Z. Gaffoor, P. Selepe, P. B. Makhoba, M. P. Mathebula, P. Mda, T. Adonis, K. S. Mapetla, B. Modibedi, T. Philip, G. Kobane, C. Bentley, S. Ramirez, S. Takuva, M. Jones, M. Sikhosana, M. Atujuna, M. Andrasik, N. S. Hejazi, A. Puren, L. Wiesner, S. Phogat, C. D. Granados, M. Koutsoukos, O. V. D. Meeren, S. W. Barnett, N. Kanesa-Thasan, J. G. Kublin, M. J. McElrath, P. B. Gilbert, H. Janes, L. Corey, Vaccine Efficacy of ALVAC-HIV and Bivalent Subtype C gp120–MF59 in Adults. *N. Engl. J. Med.* **384**, 1089–1100 (2021).

117. D. H. Barouch, F. L. Tomaka, F. Wegmann, D. J. Stieh, G. Alter, M. L. Robb, N. L. Michael, L. Peter, J. P. Nkolola, E. N. Borducchi, A. Chandrashekar, D. Jetton, K. E. Stephenson, W. Li, B. Korber, G. D. Tomaras, D. C. Montefiori, G. Gray, N. Frahm, M. J. McElrath, L. Baden, J. Johnson, J. Hutter, E. Swann, E. Karita, H. Kibuuka, J. Mpendo, N. Garrett, K. Mngadi, K. Chinyenze, F. Priddy, E. Lazarus, F. Laher, S. Nitayapan, P. Pitisuttithum, S. Bart, T. Campbell, R. Feldman, G. Lucksinger, C. Borremans, K. Callewaert, R. Roten, J. Sadoff, L. Scheppler, M. Weijtens, K. F. Boer, D. van Manen, J. Vreugdenhil, R. Zahn, L. Lavreys, S. Nijs, J. Tolboom, J. Hendriks, Z. Euler, M. G. Pau, H. Schuitemaker, Evaluation of a mosaic HIV-1 vaccine in a multicentre, randomised, double-blind, placebo-controlled, phase 1/2a clinical trial (APPROACH) and in rhesus monkeys (NHP 13-19). *The Lancet* **392**, 232–243 (2018).
118. J. Jardine, J.-P. Julien, S. Menis, T. Ota, O. Kalyuzhniy, A. McGuire, D. Sok, P.-S. Huang, S. MacPherson, M. Jones, T. Nieuwma, J. Mathison, D. Baker, A. B. Ward, D. R. Burton, L. Stamatatos, D. Nemazee, I. A. Wilson, W. R. Schief, Rational HIV Immunogen Design to Target Specific Germline B Cell Receptors. *Science* **340**, 711–716 (2013).
119. A. Escolano, H. B. Gristick, M. E. Abernathy, J. Merkenschlager, R. Gautam, T. Y. Oliveira, J. Pai, A. P. West, C. O. Barnes, A. A. Cohen, H. Wang, J. Golijanin, D. Yost, J. R. Keeffe, Z. Wang, P. Zhao, K.-H. Yao, J. Bauer, L. Nogueira, H. Gao, A. V. Voll, D. C. Montefiori, M. S. Seaman, A. Gazumyan, M. Silva, A. T. McGuire, L. Stamatatos, D. J. Irvine, L. Wells, M. A. Martin, P. J. Bjorkman, M. C. Nussenzweig, Immunization expands B cells specific to HIV-1 V3 glycan in mice and macaques. *Nature* **570**, 468–473 (2019).
120. A. Escolano, H. B. Gristick, R. Gautam, A. T. DeLaitch, M. E. Abernathy, Z. Yang, H. Wang, M. A. G. Hoffmann, Y. Nishimura, Z. Wang, N. Koranda, L. M. Kakutani, H. Gao, P. N. P. Gnanapragasam, H. Raina, A. Gazumyan, M. Cipolla, T. Y. Oliveira, V. Ramos, D. J. Irvine, M. Silva, A. P. West, J. R. Keeffe, C. O. Barnes, M. S. Seaman, M. C. Nussenzweig, M. A. Martin, P. J. Bjorkman, Sequential immunization of macaques elicits heterologous neutralizing antibodies targeting the V3-glycan patch of HIV-1 Env. *Sci. Transl. Med.* **13**, eabk1533 (2021).
121. D. J. Leggat, K. W. Cohen, J. R. Willis, W. J. Fulp, A. C. deCamp, O. Kalyuzhniy, C. A. Cottrell, S. Menis, G. Finak, L. Ballweber-Fleming, A. Srikanth, J. R. Plyler, T. Schiffner, A. Liguori, F. Rahaman, A. Lombardo, V. Philiponis, R. E. Whaley, A. Seese, J. Brand, A. M. Ruppel, W. Hoyland, N. L. Yates, L. D. Williams, K. Greene, H. Gao, C. R. Mahoney, M. M. Corcoran, A. Cagigi, A. Taylor, D. M. Brown, D. R. Ambrozak, T. Sincomb, X. Hu, R. Tingle, E. Georgeson, S. Eskandarzadeh, N. Alavi, D. Lu, T.-M. Mullen, M. Kubitz, B. Groschel, J. Maenza, O. Kolokythas, N. Khati, J.

- Bethony, S. Crotty, M. Roederer, G. B. Karlsson Hedestam, G. D. Tomaras, D. Montefiori, D. Diemert, R. A. Koup, D. S. Laufer, M. J. McElrath, A. B. McDermott, W. R. Schief, Vaccination induces HIV broadly neutralizing antibody precursors in humans. *Science* **378**, eadd6502 (2022).
122. K. W. Cohen, S. C. De Rosa, W. J. Fulp, A. C. deCamp, A. Fiore-Gartland, C. R. Mahoney, S. Furth, J. Donahue, R. E. Whaley, L. Ballweber-Fleming, A. Seese, K. Schwedhelm, D. Geraghty, G. Finak, S. Menis, D. J. Leggat, F. Rahaman, A. Lombardo, B. R. Borate, V. Philiponis, J. Maenza, D. Diemert, O. Kolokythas, N. Khati, J. Bethony, O. Hyrien, D. S. Laufer, R. A. Koup, A. B. McDermott, W. R. Schief, M. J. McElrath, A first-in-human germline-targeting HIV nanoparticle vaccine induced broad and publicly targeted helper T cell responses. *Sci. Transl. Med.* **15**, eadf3309 (2023).
 123. J. F. Scheid, H. Mouquet, N. Feldhahn, M. S. Seaman, K. Velinzon, J. Pietzsch, R. G. Ott, R. M. Anthony, H. Zebroski, A. Hurley, A. Phogat, B. Chakrabarti, Y. Li, M. Connors, F. Pereyra, B. D. Walker, H. Wardemann, D. Ho, R. T. Wyatt, J. R. Mascola, J. V. Ravetch, M. C. Nussenzweig, Broad diversity of neutralizing antibodies isolated from memory B cells in HIV-infected individuals. *Nature* **458**, 636–640 (2009).
 124. A. P. West, L. Scharf, J. F. Scheid, F. Klein, P. J. Bjorkman, M. C. Nussenzweig, Structural Insights on the Role of Antibodies in HIV-1 Vaccine and Therapy. *Cell* **156**, 633–648 (2014).
 125. L. E. McCoy, D. R. Burton, Identification and specificity of broadly neutralizing antibodies against HIV. *Immunol. Rev.* **275**, 11–20 (2017).
 126. P. D. Kwong, J. R. Mascola, HIV-1 Vaccines Based on Antibody Identification, B Cell Ontogeny, and Epitope Structure. *Immunity* **48**, 855–871 (2018).
 127. L. Gieselmann, A. T. DeLaitch, M. Rohde, C. Radford, J. Worczinski, A. Ashurov, E. Ahmadov, J. A. Burger, C. Havenar-Daughton, S. Deshpande, F. Giovannoni, D. Corti, C. Kreer, M. S. Ercanoglu, P. Schommers, I. S. Georgiev, A. P. West, J. Knüfer, R. Stumpf, A. Kroidl, C. Geldmacher, L. Maganga, W. William, N. E. Ntinginya, M. Hoelscher, Z. Yang, Q. Wei, M. B. Renfrow, T. J. Green, J. Novak, M. J. van Gils, H. B. Gristick, H. Gruell, J. D. Bloom, M. S. Seaman, P. J. Bjorkman, F. Klein, Identification of a potent V3 glycan site broadly neutralizing antibody targeting an N332gp120 glycan-independent epitope. *Nat. Immunol.* **27**, 572–585 (2026).
 128. J. R. Mascola, M. G. Lewis, G. Stiegler, D. Harris, T. C. VanCott, D. Hayes, M. K. Louder, C. R. Brown, C. V. Sapan, S. S. Frankel, Y. Lu, M. L. Robb, H. Katinger, D.

- L. Birx, Protection of Macaques against Pathogenic Simian/Human Immunodeficiency Virus 89.6PD by Passive Transfer of Neutralizing Antibodies. *J. Virol.* **73**, 4009–4018 (1999).
129. J. R. Mascola, G. Stiegler, T. C. VanCott, H. Katinger, C. B. Carpenter, C. E. Hanson, H. Beary, D. Hayes, S. S. Frankel, D. L. Birx, M. G. Lewis, Protection of macaques against vaginal transmission of a pathogenic HIV-1/SIV chimeric virus by passive infusion of neutralizing antibodies. *Nat. Med.* **6**, 207–210 (2000).
 130. S.-Y. Ko, A. Pegu, R. S. Rudicell, Z. Yang, M. G. Joyce, X. Chen, K. Wang, S. Bao, T. D. Kraemer, T. Rath, M. Zeng, S. D. Schmidt, J.-P. Todd, S. R. Penzak, K. O. Saunders, M. C. Nason, A. T. Haase, S. S. Rao, R. S. Blumberg, J. R. Mascola, G. J. Nabel, Enhanced neonatal Fc receptor function improves protection against primate SHIV infection. *Nature* **514**, 642–645 (2014).
 131. M. Caskey, F. Klein, J. C. C. Lorenzi, M. S. Seaman, A. P. West, N. Buckley, G. Kremer, L. Nogueira, M. Braunschweig, J. F. Scheid, J. A. Horwitz, I. Shimeliovich, S. Ben-Avraham, M. Witmer-Pack, M. Platten, C. Lehmann, L. A. Burke, T. Hawthorne, R. J. Gorelick, B. D. Walker, T. Keler, R. M. Gulick, G. Fätkenheuer, S. J. Schlesinger, M. C. Nussenzweig, Viraemia suppressed in HIV-1-infected humans by broadly neutralizing antibody 3BNC117. *Nature* **522**, 487–491 (2015).
 132. M. Caskey, T. Schoofs, H. Gruell, A. Settler, T. Karagounis, E. F. Kreider, B. Murrell, N. Pfeifer, L. Nogueira, T. Y. Oliveira, G. H. Learn, Y. Z. Cohen, C. Lehmann, D. Gillor, I. Shimeliovich, C. Unson-O'Brien, D. Weiland, A. Robles, T. Kümmerle, C. Wyen, R. Levin, M. Witmer-Pack, K. Eren, C. Ignacio, S. Kiss, A. P. West, H. Mouquet, B. S. Zingman, R. M. Gulick, T. Keler, P. J. Bjorkman, M. S. Seaman, B. H. Hahn, G. Fätkenheuer, S. J. Schlesinger, M. C. Nussenzweig, F. Klein, Antibody 10-1074 suppresses viremia in HIV-1-infected individuals. *Nat. Med.* **23**, 185–191 (2017).
 133. C. Gaebler, L. Nogueira, E. Stoffel, T. Y. Oliveira, G. Breton, K. G. Millard, M. Turroja, A. Butler, V. Ramos, M. S. Seaman, J. D. Reeves, C. J. Petropoulos, I. Shimeliovich, A. Gazumyan, C. S. Jiang, N. Jilg, J. F. Scheid, R. Gandhi, B. D. Walker, M. C. Sneller, A. Fauci, T.-W. Chun, M. Caskey, M. C. Nussenzweig, Prolonged viral suppression with anti-HIV-1 antibody therapy. *Nature* **606**, 368–374 (2022).
 134. M. C. Sneller, J. Blazkova, J. S. Justement, V. Shi, B. D. Kennedy, K. Gittens, J. Tolstenko, G. McCormack, E. J. Whitehead, R. F. Schneck, M. A. Proschan, E. Benko, C. Kovacs, C. Oguz, M. S. Seaman, M. Caskey, M. C. Nussenzweig, A. S. Fauci, S. Moir, T.-W. Chun, Combination anti-HIV antibodies provide sustained virological suppression. *Nature* **606**, 375–381 (2022).

135. L. Corey, P. B. Gilbert, M. Juraska, D. C. Montefiori, L. Morris, S. T. Karuna, S. Edupuganti, N. M. Mgodl, A. C. deCamp, E. Rudnicki, Y. Huang, P. Gonzales, R. Cabello, C. Orrell, J. R. Lama, F. Laher, E. M. Lazarus, J. Sanchez, I. Frank, J. Hinojosa, M. E. Sobieszczyk, K. E. Marshall, P. G. Mukwekwerere, J. Makhema, L. R. Baden, J. I. Mullins, C. Williamson, J. Hural, M. J. McElrath, C. Bentley, S. Takuva, M. M. G. Lorenzo, D. N. Burns, N. Espy, A. K. Randhawa, N. Kochar, E. Piwowar-Manning, D. J. Donnell, N. Sista, P. Andrew, J. G. Kublin, G. Gray, J. E. Ledgerwood, J. R. Mascola, M. S. Cohen, Two Randomized Trials of Neutralizing Antibodies to Prevent HIV-1 Acquisition. *N. Engl. J. Med.* **384**, 1003–1014 (2021).
136. C. H. Fellingner, M. R. Gardner, J. A. Weber, B. Alfant, A. S. Zhou, M. Farzan, eCD4-Ig Limits HIV-1 Escape More Effectively than CD4-Ig or a Broadly Neutralizing Antibody. *J. Virol.* **93**, 10.1128/jvi.00443-19 (2019).
137. D. J. Capon, S. M. Chamow, J. Mordenti, S. A. Marsters, T. Gregory, H. Mitsuya, R. A. Byrn, C. Lucas, F. M. Wurm, J. E. Groopman, S. Broder, D. H. Smith, Designing CD4 immunoadhesins for AIDS therapy. *Nature* **337**, 525–531 (1989).
138. J. M. Jacobson, I. Lowy, C. V. Fletcher, T. J. O'Neill, D. N. H. Tran, T. J. Ketas, A. Trkola, M. E. Klotman, P. J. Maddon, W. C. Olson, R. J. Israel, Single-Dose Safety, Pharmacology, and Antiviral Activity of the Human Immunodeficiency Virus (HIV) Type 1 Entry Inhibitor PRO 542 in HIV-Infected Adults. *J. Infect. Dis.* **182**, 326–329 (2000).
139. W. T. Shearer, R. J. Israel, S. Starr, C. V. Fletcher, D. Wara, M. Rathore, J. Church, J. DeVille, T. Fenton, B. Graham, P. Samson, S. Staprans, J. McNamara, J. Moye, P. J. Maddon, W. C. Olson, for the Pediatric AIDS Clinical Trials Group Protocol 351 Study Team, Recombinant CD4-IgG2 in Human Immunodeficiency Virus Type 1—Infected Children: Phase 1/2 Study. *J. Infect. Dis.* **182**, 1774–1779 (2000).
140. C. V. Fletcher, J. G. DeVille, P. M. Samson, J. H. Moye, J. A. Church, H. M. L. Spiegel, P. Palumbo, T. Fenton, M. E. Smith, B. Graham, J. M. Kraisner, W. T. Shearer, Nonlinear pharmacokinetics of high-dose recombinant fusion protein CD4-IgG2 (PRO 542) observed in HIV-1-infected children. *J. Allergy Clin. Immunol.* **119**, 747–750 (2007).
141. A. J. Bahn-Suh, L. F. Caldera, P. N. P. Gnanapragasam, J. R. Keeffe, M. S. Seaman, P. J. Bjorkman, S. L. Mayo, Stabilized gp120-specific CD4 for next-generation HIV-1 inhibitors. bioRxiv [Preprint] (2026). <https://doi.org/10.64898/2026.03.24.713825>.

142. C. Doyle, J. L. Strominger, Interaction between CD4 and class II MHC molecules mediates cell adhesion. *Nature* **330**, 256–259 (1987).
143. W. Chen, Y. Feng, P. Prabakaran, T. Ying, Y. Wang, J. Sun, C. D. S. Macedo, Z. Zhu, Y. He, V. R. Polonis, D. S. Dimitrov, Exceptionally Potent and Broadly Cross-Reactive, Bispecific Multivalent HIV-1 Inhibitors Based on Single Human CD4 and Antibody Domains. *J. Virol.* **88**, 1125–1139 (2014).
144. M. Thali, J. P. Moore, C. Furman, M. Charles, D. D. Ho, J. Robinson, J. Sodroski, Characterization of conserved human immunodeficiency virus type 1 gp120 neutralization epitopes exposed upon gp120-CD4 binding. *J. Virol.* **67**, 3978–3988 (1993).
145. S.-H. Xiang, L. Wang, M. Abreu, C.-C. Huang, P. D. Kwong, E. Rosenberg, J. E. Robinson, J. Sodroski, Epitope mapping and characterization of a novel CD4-induced human monoclonal antibody capable of neutralizing primary HIV-1 strains. *Virology* **315**, 124–134 (2003).
146. R. Diskin, P. M. Marcovecchio, P. J. Bjorkman, Structure of a clade C HIV-1 gp120 bound to CD4 and CD4-induced antibody reveals anti-CD4 polyreactivity. *Nat. Struct. Mol. Biol.* **17**, 608–613 (2010).
147. A. P. West, R. P. Galimidi, C. P. Foglesong, P. N. P. Gnanapragasam, J. S. Klein, P. J. Bjorkman, Evaluation of CD4-CD4i Antibody Architectures Yields Potent, Broadly Cross-Reactive Anti-Human Immunodeficiency Virus Reagents. *J. Virol.* **84**, 261–269 (2010).
148. H. Wang, C. O. Barnes, Z. Yang, M. C. Nussenzweig, P. J. Bjorkman, Partially Open HIV-1 Envelope Structures Exhibit Conformational Changes Relevant for Coreceptor Binding and Fusion. *Cell Host Microbe* **24**, 579-592.e4 (2018).
149. J. Charles A Janeway, P. Travers, M. Walport, M. J. Shlomchik, “The structure of a typical antibody molecule” in *Immunobiology: The Immune System in Health and Disease*. 5th Edition (Garland Science, 2001; <https://www.ncbi.nlm.nih.gov/books/NBK27144/>).
150. P. Holliger, P. J. Hudson, Engineered antibody fragments and the rise of single domains. *Nat. Biotechnol.* **23**, 1126–1136 (2005).
151. S. Muyldermans, Nanobodies: Natural Single-Domain Antibodies. *Annu. Rev. Biochem.* **82**, 775–797 (2013).

152. C. McMahon, A. S. Baier, R. Pascolutti, M. Wegrecki, S. Zheng, J. X. Ong, S. C. Erlandson, D. Hilger, S. G. F. Rasmussen, A. M. Ring, A. Manglik, A. C. Kruse, Yeast surface display platform for rapid discovery of conformationally selective nanobodies. *Nat. Struct. Mol. Biol.* **25**, 289–296 (2018).

DESIGNED MOSAIC NANOPARTICLES ENHANCE CROSS-REACTIVE IMMUNE RESPONSES IN MICE

E. Wang*, A. A. Cohen*, **L. F. Caldera***, J. R. Keeffe, A. V. Rorick, Y. M. Adia, P. N. P. Gnanapragasam, P. J. Bjorkman, A. K. Chakraborty, Designed mosaic nanoparticles enhance cross-reactive immune responses in mice. *Cell* 188, 1036-1050.e11 (2025).
<https://doi.org/10.1016/j.cell.2024.12.015>

** These authors contributed equally.*

2.1 Summary

Nanoparticle vaccines displaying combinations of SARS-like betacoronavirus (sarbecovirus) receptor-binding domains (RBDs) could protect against SARS-CoV-2 variants and spillover of zoonotic sarbecoviruses into humans. Using a computational approach, we designed variants of SARS-CoV-2 RBDs and selected 7 natural sarbecovirus RBDs, each predicted to fold properly and abrogate antibody responses to variable epitopes. RBDs were attached to 60-mer nanoparticles to make immunogens displaying two (mosaic-2_{COMS}), five (mosaic-5_{COM}), or seven (mosaic-7_{COM}) different RBDs for comparisons with mosaic-8b, which elicited cross-reactive antibodies and protected animals from sarbecovirus challenges. Naive and COVID-19 pre-vaccinated mice immunized with mosaic-7_{COM} elicited antibodies targeting conserved RBD epitopes, and their sera exhibited higher binding and neutralization titers against sarbecoviruses than mosaic-8b. Mosaic-2_{COMS} and mosaic-5_{COM} elicited higher antibody potencies against some SARS-CoV-2 variants than mosaic-7_{COM}. However, mosaic-7_{COM} elicited more potent responses against zoonotic sarbecoviruses and highly mutated Omicrons, supporting its use to protect against SARS-CoV-2 variants and zoonotic sarbecoviruses.

2.2 Introduction

Emerging SARS-CoV-2 variants, notably Omicron and its subvariants, have demonstrated the ability to partially evade previous vaccine-induced immune responses by mutating epitopes targeted by the generated antibodies, (1–9) thus extending the duration and impact of the COVID-19 pandemic. While mRNA vaccines have been adapted to include sequences based on existing Omicron strains, they become outdated due to the continuous emergence of new variants (10). Moreover, there remains the continuing risk of future pandemics due to cross-species spillovers from the pool of existing zoonotic SARS-like betacoronaviruses (sarbecoviruses) (11, 12). Consequently, the development of universal sarbecovirus vaccines capable of safeguarding against future SARS-CoV-2 variants and new viruses derived from sarbecoviruses is critical for public health. The aim of such universal sarbecovirus vaccines is to generate antibodies targeting conserved epitopes on spike trimers, which are not generated in high titers by mRNA vaccines encoding spike trimers for reasons outlined below.

SARS-CoV-2 uses its spike trimer to infiltrate host cells by binding to the host receptor known as angiotensin-converting enzyme 2 (ACE2) (13, 14). Specifically, the receptor-binding domain (RBD) of the spike binds to ACE2, and it can do so only when the RBD adopts an “up” conformation rather than its usual “down” conformation (15). Upon infection or vaccination with the spike trimer, numerous antibodies targeting the RBD are elicited, categorized into four primary types (classes 1, 2, 3, and 4) based on their epitopes (**Figure 1A**) (15). The epitopes of class 1 and 2 antibodies typically overlap with the ACE2 binding site on the RBD and have evolved due to immune pressure over time, while class 3 and 4 antibodies bind to more conserved but less accessible (in the case of class 4) epitopes. Notably, class 4 antibodies are sterically occluded even on “up” RBDs, making them challenging to induce by viral infection or using vaccines containing spike trimers, as shown by deep mutation scanning (DMS) mapping of antisera from convalescent COVID-19 or vaccinated donors (16–20). A vaccine capable of eliciting antibodies against the class 4 and class 1/4 (class 4-like antibodies that reach toward the class 1 epitope and sterically occlude

ACE2 binding) epitopes (21) could target conserved sites, providing protection against future SARS-CoV-2 variants and potential sarbecovirus spillovers.

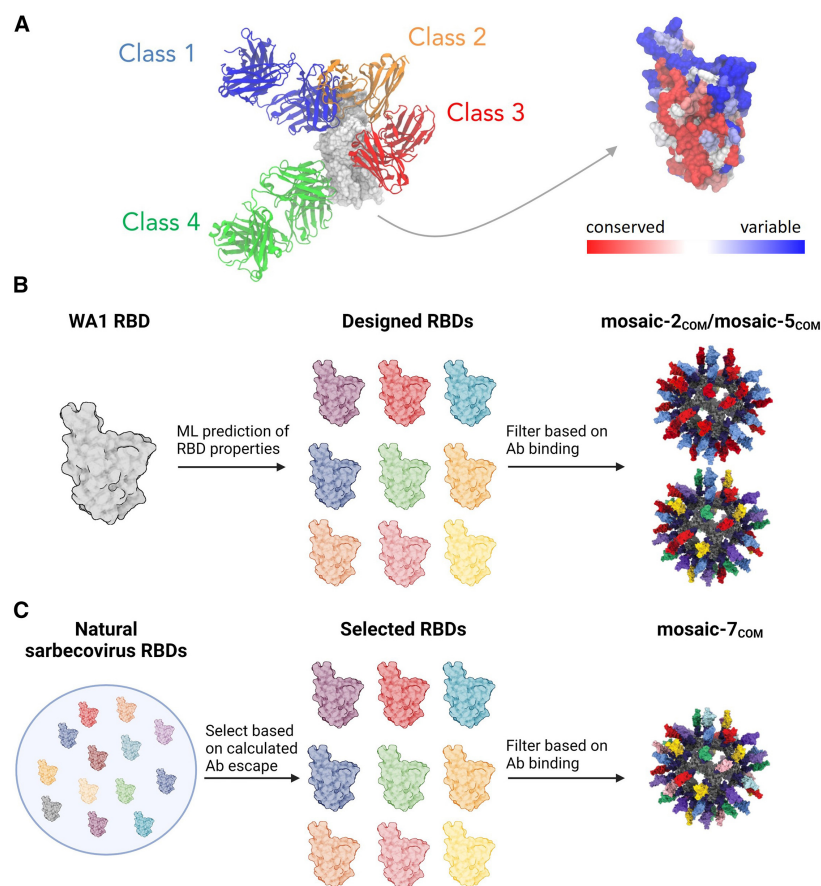


Figure 1. Overview of the design process

(A) Structures of representative class 1 (C102, PDB: 7K8M), class 2 (C144, PDB: 7K90), class 3 (S309, PDB: 7JMX), and class 4 (CR3022, PDB: 6W41) antibodies bound to the WA1 SARS-CoV-2 RBD, and the structure of the WA1 RBD (PDB: 6W41) colored based on conservation scores calculated using the ConSurf database (22).

(B) Overview of mosaic-2_{COM} and mosaic-5_{COM} RBD-NP designs. Starting from the WA1 RBD, computational analysis and machine learning models (23) were used to calculate properties of potential RBD immunogens based on expression, antibody binding, and solubility. A set of selected RBDs was further filtered based on expression and binding measurements and used to construct the mosaic-2_{COM} and mosaic-5_{COM} RBD NPs.

(C) Overview of designing mosaic-7_{COM}. A set of 8 RBDs was selected from naturally occurring zoonotic sarbecovirus RBDs to maximize (1) sequence diversity and (2) binding to class 3 and 4 but not class 1 and 2 RBD epitopes (RBD epitopes defined as described) (15). The 8 selected RBDs were further filtered based on experimentally determined properties (see text), and the 7 remaining RBDs were used for mosaic-7_{COM}.

Previously, mosaic-8b RBD nanoparticles (RBD NPs) were developed as a potential pan-sarbecovirus vaccine by using the SpyCatcher-SpyTag system (24, 25) to covalently attach different RBDs with C-terminal SpyTag sequences to a 60-mer mi3 protein NP with N-terminal SpyCatcher proteins in each subunit (26). These NPs, which displayed RBDs from SARS-CoV-2 and seven zoonotic sarbecoviruses, were hypothesized to promote the development of cross-reactive antibodies by exposing epitopes that were conserved across the RBDs. B cells displaying cross-reactive B cell receptors would be able to bind bivalently to adjacent conserved regions on the displayed RBDs (27). By contrast, B cells displaying strain-specific B cell receptors that target variable epitopes would rarely bind bivalently to mosaic RBD NPs because of their inability to bind non-identical adjacent RBDs. Therefore, the cross-reactive B cells would have an avidity advantage that would enable them to outcompete the strain-specific B cells during germinal center (GC) reactions. Animal studies supported this hypothesis, as the mosaic-8 RBD NPs elicited high titers of cross-reactive antibodies (26) and protected K18-hACE2 transgenic mice (28) and non-human primates against sarbecovirus challenges (27). Although mosaic-8b displayed a particular set of RBDs that were selected based on phylogenetics and pandemic potential, the SpyCatcher-SpyTag system is flexible and allows various combinations of proteins to be easily attached covalently in various combinations to a SpyCatcher NP. This leaves open the intriguing possibility that the displayed RBD sequences could be further optimized to generate NPs that elicit even more potent cross-reactive antibodies.

In this work, we combined computational and experimental approaches to design and test sets of new mosaic RBD NPs with the goal of eliciting improved cross-reactive responses. The first set contained RBDs designed with six mutations relative to the SARS-CoV-2 WA1

strain aimed at maintaining expression and solubility while selectively abrogating antibody binding to class 1 and class 2 RBD epitopes (**Figure 1B**). The second set contained sarbecovirus RBDs that selectively abrogated class 1 and 2 antibody binding and had the highest sequence diversity among all computationally generated sets (**Figure 1C**). After experimentally filtering the RBDs for expression, solubility, and antibody binding, we constructed mosaic RBD NPs and evaluated them in mice. Binding and pseudovirus neutralization titers from naive mice immunized with RBD NPs show that our designed RBD NPs elicited more cross-reactive responses than mosaic-8b and homotypic SARS-CoV-2 Beta RBD NPs. Deep mutational scanning profiles suggested that the antibody response is focused on class 3 and 4 RBD epitopes for the mosaic-7_{COM} RBD-NP. Finally, serum responses of mice with prior COVID-19 vaccinations showed that mosaic-7_{COM} elicited higher neutralization titers against a range of viral strains compared with mosaic-8b, mosaic-7 (mosaic-8b without SARS-CoV-2 Beta), and the bivalent WA1/BA.5 mRNA lipid nanoparticle (LNP) vaccine. Taken together, these results suggest that computationally designed RBD NPs can outperform more empirical designs and lead to promising candidates for potential pan-sarbecovirus vaccines such as mosaic-7_{COM}.

2.3 Results

2.3.1 WA1 RBDs were designed to elicit antibodies against less mutated SARS-CoV-2 variants

Our first set of RBD NPs displayed WA1 RBDs with mutations that were designed to promote the generation of cross-reactive antibodies that target relatively conserved epitopes on the RBDs of SARS-CoV-2 variants. Overall, our computational design strategy sought to create RBDs that (1) abrogated binding of class 1 and class 2 anti-RBD antibodies but not class 3, 4, and 1/4 antibodies (RBD epitopes defined as described) (15); (2) were stable and expressed well; and (3) yielded soluble RBD NPs upon conjugation.

We designed sets of two RBDs to be displayed on a particular NP, with each RBD containing 6 mutations. Although it might be ideal to design RBD NPs with more variant RBDs, with each containing numerous mutations, introducing many mutations could result in improperly

folded RBDs. Our choice of 6 mutations per RBD was informed by our method of predicting relative expression of different RBDs, which is a convolutional neural network trained on DMS experiments using a library of different RBDs displayed on yeast (29). In DMS experiments used to train our neural network, yeast cells displayed RBDs containing random mutations relative to the WA1 strain, and the expression of each variant was measured (29). The DMS-generated RBD variants contained between 0 and 7 mutations, so a model trained on these data would not be effective at predicting the expression of variants containing more than 7 mutations. We chose 6 mutations per RBD because this number is below the maximum of 7 mutations and because it is even (we divide the 6 mutations into 3 class 1 escape mutations and 3 class 2 escape mutations).

Previous DMS experiments (17, 30–33) quantified escape from antibodies (either polyclonal serum antibodies or monoclonal antibodies [mAbs]) in the following way: yeast cells for each RBD mutation were created and sorted into an antibody escape bin based on it not binding to a particular antibody or antiserum. The escape fraction of a RBD mutation is the fraction of yeast cells expressing the mutation that were in the escape bin. An escape fraction of 0 meant that none of the yeast cells expressing the mutation were in the escape bin, while a fraction of 1 meant all yeast cells expressing the mutation were in the bin.

We first considered two RBDs per NP for the following reasons. DMS data on antibody escape at the time we designed the RBD NPs were evaluated relative to the SARS-CoV-2 WA1 RBD (17, 30–33), so it was easiest to design a new RBD with mutations that abrogated binding of WA1-specific class 1 and 2 anti-RBD antibodies. However, new class 1 and 2 anti-RBD antibodies will evolve upon immunization with RBDs that abrogate binding to the usually immunodominant antibodies (34, 35). Our solution to this problem was to place escape mutations for different RBDs in different positions relative to each other, as the new germline and GC B cells that recognize class 1 and 2 epitopes on one designed RBD would likely not bind bivalently to a second RBD that contains escape mutations in different residues and would thus be at a disadvantage compared with the class 3, 4, and 1/4 antibodies. Based on this hypothesis, one would ideally create an RBD-NP with many different RBDs

if their escape mutations were all in different positions. However, we were limited to 6 mutations for the reasons stated above, so we decided to use 2 RBDs per NP to introduce more escape mutations in each RBD and therefore increase the probability of abrogating bivalent antibody binding. However, using only 2 RBDs per NP to create mosaic-2 RBD NPs results in a higher probability that neighboring RBDs are identical: the average probability of neighboring identical RBDs in a mosaic-2 is 0.5, whereas the average probability of neighboring identical RBDs in a mosaic-8 is 0.125. We also created a mosaic-5_{COM} RBD-NP to empirically determine whether displaying more variant RBDs with some shared mutations would result in differences in cross-reactive antibody elicitation.

First, we determined the 20 RBD positions with the highest escapes from class 1 and 2 anti-RBD antibodies based on DMS data (17, 30–33) (Tables S1 and S2). We chose to focus on 20 RBD positions because a previous DMS study highlighted ~20 positions where mutations affected binding to class 1 and 2 anti-RBD antibodies (31). Mutations in these 20 positions do not necessarily occur all at once, e.g., the BA.1 SARS-CoV-2 Omicron variant contains substitutions in 15 RBD positions relative to the WA1 RBD. From these 20 positions, we generated $\binom{20}{6} = 38,760$ combinations of 6 positions for both class 1 and class 2 escape positions. Because we have 2 RBDs on a mosaic-2_{COM}, we divided each combination of 6 positions into 2 groups of 3, for which there are 10 possible enumerations, generating 387,600 sets for class 1 RBD positions and 387,600 sets for class 2 positions, as shown in **Figure S1**. Creating all possible 387,600² RBD pairs by combining the class 1 and class 2 sets is computationally infeasible, so we instead randomly sampled ~800,000 RBD pairs for further evaluation.

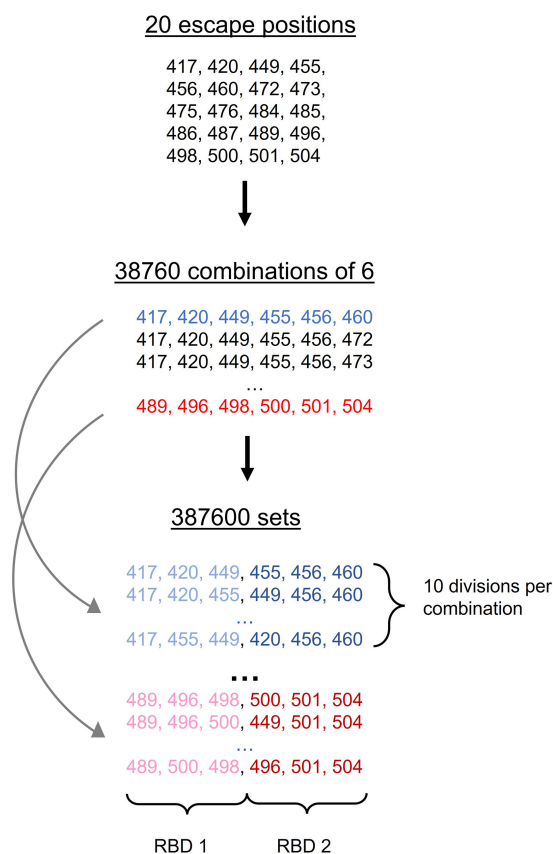


Figure S1. Illustration of how sequences for class 1 escape mutations were generated, related to Figure 2

From the 20 RBD positions with the highest escapes against class 1 anti-RBD antibodies, we generated all 38,760 possible combinations of 6 positions. For each combination, we further generated all possible ways to divide the 6 positions into 2 groups of 3, of which there were 10 possible divisions, resulting in 387,600 sets of positions. For each set, 1 group of 3 positions was assigned to RBD1 (colored as light blue or pink), and the other group of 3 positions was assigned to RBD2 (colored as dark blue or red). RBD1 and RBD2 would be mutated in their assigned positions to the amino acids listed in Table S2.

Additionally, for a particular escape position, the amino acid mutation with the largest escape fraction that was also not a charged-to-hydrophobic substitution was chosen (Table S2). The decision to avoid charged-to-hydrophobic substitutions was meant to enhance solubility, as preliminary RBD designs showed aggregation when charged-to-hydrophobic mutations were included. For example, RBD residue 484 is a class 2 escape residue (31), and the

corresponding escape mutation we choose was E484R, which exhibited the largest escape fraction for non-hydrophobic amino acids (17, 30–33).

The ~800,000 pairs of RBD sequences were then screened for likelihood of successful expression using a convolutional neural network that was previously trained on DMS data (23) (Figures 2A and 2B). We selected RBD pairs for which both RBDs were predicted to express well, which was defined as having a change in expression from WA1 greater than -0.2 log-mean fluorescence intensity (logMFI) based on DMS data (29). This threshold was previously chosen such that sequences of circulating variants, which are known to express well because they are found in nature, had predicted logMFI values above this threshold (29). Of the ~800,000 RBD pairs, ~100,000 were selected that fit the chosen computational expression criterion.

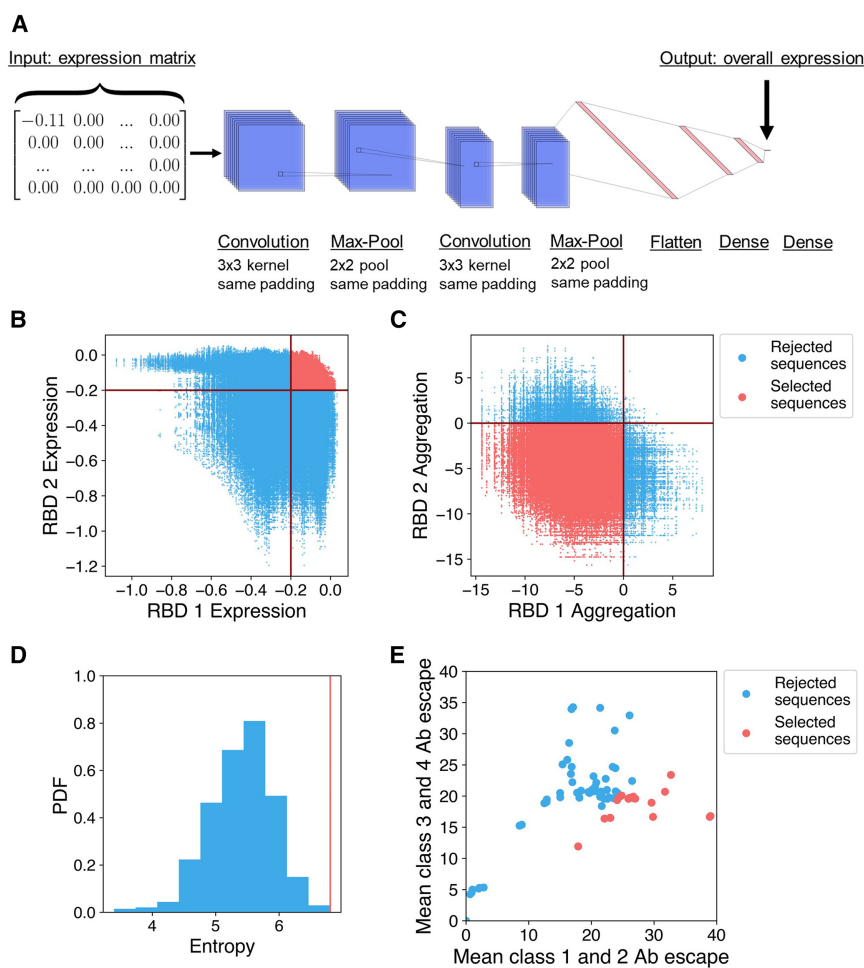


Figure 2. Overview of computational methods

(A) Architecture of the neural network used to predict RBD expression (23). The input is an expression matrix, which is the element-wise product (multiplication of entries at the same positions) of the one-hot encoded sequence (each residue is represented as a 20D vector with entries of 1 for the matching amino acid and 0 for other amino acids) and the matrix of single-mutation expression changes. This is processed through a convolutional neural network to produce the predicted change in expression as an output.

(B) ~800,000 possible RBD sequences are screened for predicted expression relative to the WA1 RBD using a threshold value of $-0.2 \log\text{MFI}$. Rejected RBD pairs are in blue, and selected pairs are in red.

(C) ~100,000 RBD sequences that passed predicted expression screening and were further screened for solubility based on a change in aggregation score relative to WA1 calculated using Aggrescan. Rejected RBD pairs are in blue, and selected pairs are in red.

(D) The distribution of total mutational entropy over sets of 10 RBDs, and the set selected for experimental testing is the one with maximum entropy indicated by the red line.

(E) Mean escape against class 1 and 2 anti-RBD antibodies and the mean escape against class 3 and 4 anti-RBD antibodies for naturally occurring sarbecoviruses. Rejected RBDs are in blue, and selected RBDs are in red.

The ~100,000 selected pairs of RBD sequences were further evaluated for predicted solubility using Aggrescan (36) to calculate the aggregation score of each RBD in the pair relative to the WA1 RBD (**Figure 2C**). We selected ~90,000 RBD pairs for which both RBDs were predicted to be more soluble than the WA1 RBD. The large fraction (~0.9) of selected pairs suggests that the avoidance of charged-to-hydrophobic mutations in previous steps was effective at preserving predicted solubility.

Of these ~90,000 RBD pairs, we selected the top 20,000 in terms of total class 1 and class 2 antibody escape (estimated as a sum over the escape fractions for mutated residues on both RBDs) to further reduce recognition of class 1 and class 2 RBD epitopes. We selected 20,000 because the total class 1 and class 2 antibody escape plateaus after the top 20,000 pairs (**Figure S2**). From these 20,000, we selected a subset for experimental testing. We

computationally designed multiple RBD pairs in case some RBDs failed to express, abrogate antibody binding, or remain soluble with limited aggregation. In creating these RBD pairs, we sought to avoid pairs that were very similar. Therefore, we randomly selected sets of 5 RBD pairs, calculated the total mutational entropy, and selected the set with the highest entropy (**Figure 2D**). More specifically, each set of 5 RBD pairs contained 10 RBDs, and we calculated the Shannon entropy (37) for each residue over the 10 RBDs. The total mutational entropy was then the sum of the Shannon entropies for all residues. The RBD sequences are reported in Table S3.

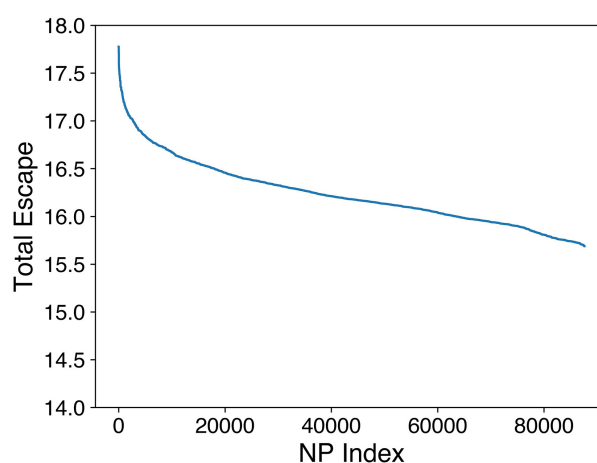


Figure S2. The total escape against class 1 and 2 antibodies for all ~90,000 RBD pairs that pass screening, related to Figure 2

The total escape is obtained from DMS experiments (17, 30–33) and antibody RBD-epitope classes are defined in Barnes et al. (15)

2.3.2 Zoonotic sarbecovirus RBDs were selected to elicit cross-reactive antibodies against sarbecoviruses

Additionally, we selected RBDs from various sarbecoviruses to make a new mosaic RBD-NP in a manner distinct from choices for mosaic-8b RBD-NP (27). While mosaic-8b used phylogenetics and pandemic potential in its design (selecting clade 1, clade 1b, and clade 2 sarbecovirus RBDs from a study of RBD receptor usage and cell tropism (38)), we instead used antibody binding data to select RBDs. We first obtained a set of 246 non-redundant sarbecovirus RBDs from the NCBI database (39), aligned these with the WA1 SARS-CoV-

2 RBD using ClustalW (40, 41), and filtered the alignment for residues 331–531 of the WA1 SARS-CoV-2 spike since these were the WA1 spike residues used for RBD display in DMS experiments (17, 30–33). For each RBD in the alignment, we examined its substitutions relative to the WA1 RBD amino acids and calculated the substitutions' average escapes to class 1, 2, 3, and 4 anti-RBD antibodies from the DMS data. The selective binding B_s of each RBD was then scored using

$$B_s = \langle B_1 \rangle + \langle B_2 \rangle - \langle B_3 \rangle - \langle B_4 \rangle \quad \text{Equation 1}$$

Where $\langle B_i \rangle$ is the average total escape of an RBD to antibodies of class i . Thus, RBDs that have high escapes from class 1 and 2 antibodies but low escapes from class 3 and 4 antibodies would maximize B_s . We selected the top 40 sarbecovirus RBDs in terms of B_s . In **Figure 2E**, we graph the mean class 1 and 2 escapes ($\langle B_1 \rangle + \langle B_2 \rangle$) and mean class 3 and 4 escapes ($\langle B_3 \rangle + \langle B_4 \rangle$) for every sarbecovirus RBD and highlight the selected RBDs in red. The selected RBDs clustered toward the lower right region, demonstrating that our calculation of B_s selected for high class 1 and 2 escapes and low class 3 and 4 escapes. From the selected RBDs, we generated sets of 8 as in previous studies (26, 27). We then calculated the fraction of amino acids that were the same for every pair of RBDs (average pairwise amino acid sequence identity as defined to create mosaic-8b). We selected the set of 8 with the lowest average amino acid sequence identity between pairs (Table S4).

2.3.3 Designed RBDs bind class 3 and 4 anti-RBD antibodies and conjugate to form stable RBD NPs

Before creating mosaic RBD NPs with the computationally designed RBDs, we experimentally evaluated their expression and binding to characterized anti-RBD mAbs, removing any candidates that showed suboptimal properties. First, we expressed the RBDs and purified them from transfected cell supernatants using nickel-nitrilotriacetic acid (Ni-NTA) affinity chromatography followed by size exclusion chromatography (SEC) (**Figures 3A and 3B**). For the designed RBDs, 8 of 10 exhibited expected high levels of expression, while one expressed at low levels (RBD8) and another showed no detectable expression

based on SEC chromatograms (RBD3) (**Figure 3A**). RBD3 and RBD8 were therefore removed from further consideration. Given that 70% of single RBD substitutions eliminated expression in a DMS library (29), generating 6-mutant RBDs that preserve expression with an 80% success rate is notably efficient and points to the utility of our neural network predictor. All zoonotic sarbecovirus RBDs expressed effectively (**Figure 3B**) as expected because well-folded RBDs are likely to be found in nature.

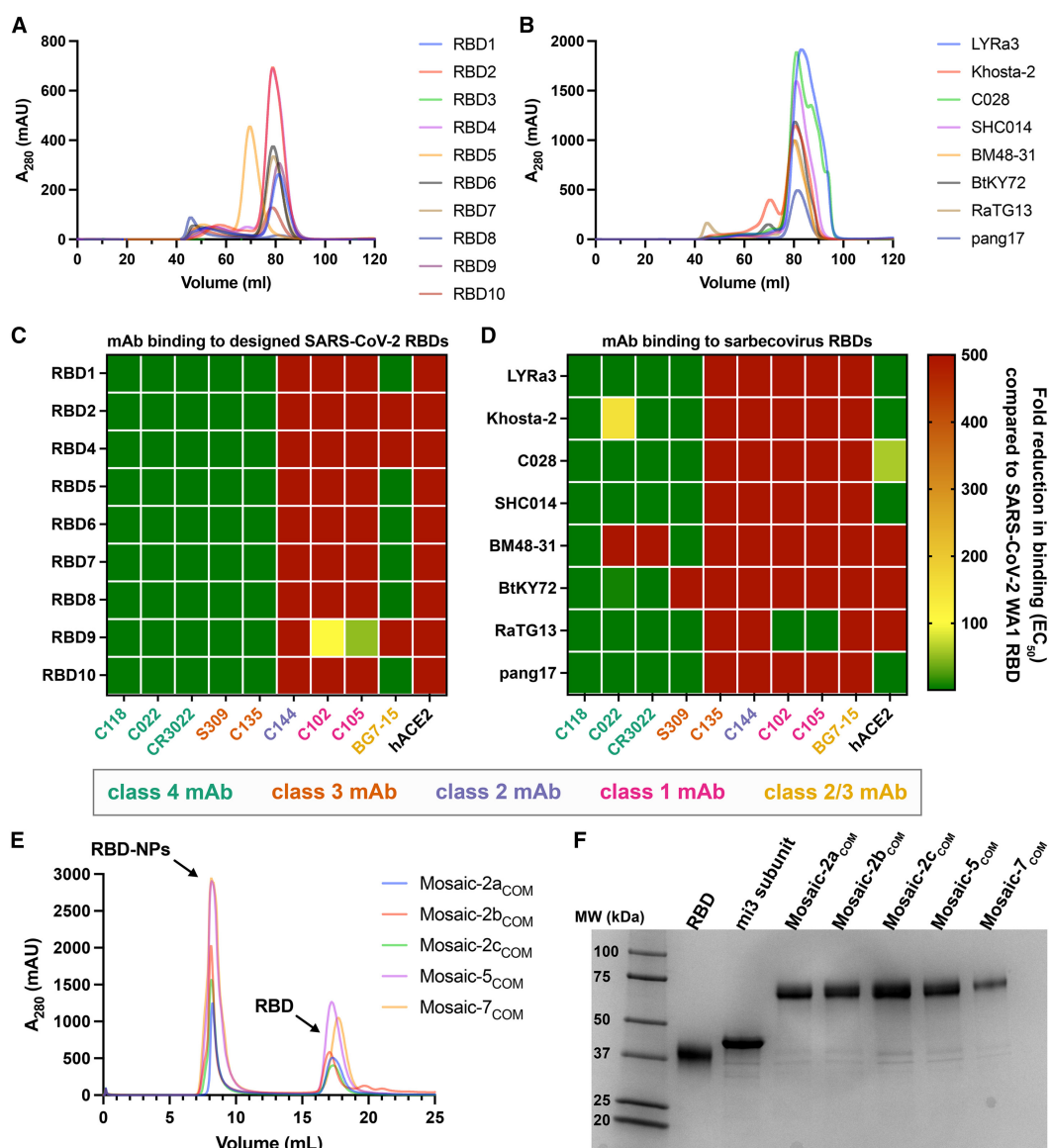


Figure 3. Designed SARS-CoV-2 RBDs and sarbecovirus RBDs exhibit desired properties

(A) and (B) HiLoad 16/600 Superdex 200 SEC profiles of designed RBDs (A) and sarbecovirus RBDs (B). RBD3 and RBD8 exhibited suboptimal expression, indicated by no signal for an RBD monomer (RBD3) or a peak in the void volume (RBD8).

(C) and (D) Fold reduction of selected monoclonal anti-RBD antibodies (mAbs) or a human ACE2-Fc construct (hACE2) to designed SARS-CoV-2 RBDs (C) and sarbecovirus RBDs (D) compared with binding to WA1 RBD.

(E) Superose 6 increase 10/300 SEC profiles after SpyTagged RBDs were conjugated to SpyCatcher-mi3 showing peaks for RBD NPs and free RBDs.

(F) SDS-PAGE for each RBD-NP after pooling appropriate SEC fractions.

We then used ELISAs to derive binding EC_{50} s of these RBDs to a panel of mAbs directed against class 1, 2, 3, and 4 RBD epitopes, demonstrating that the designed RBDs bound class 3 and 4, but not class 1 or class 2, anti-RBD antibodies (**Figure 3C**). Interestingly, the class 2/3 antibody BG7-15 (42) exhibited mixed results, binding to RBD1, RBD5, RBD6, RBD7, RBD8, and RBD10 but not to RBD2, RBD4, or RBD9 (**Figure 3C**). Although inconsequential for immunization purposes, none of the designed RBDs bound to a human ACE2-Fc construct because class 1 and 2 escape mutations are located near the ACE2 binding site (15, 17, 30–33). The EC_{50} s of the zoonotic sarbecovirus RBDs for binding the panel of antibodies showed the same trends, although some RBDs (Khosta-2, BM48-31, BtKY72) did not bind all class 3 or 4 antibodies (**Figure 3D**). RaTG13 RBD retained binding to class 1 antibodies, so it was removed from consideration.

From the designed RBDs, we created 3 RBD NPs displaying 2 RBDs each (mosaic-2_{acom}, mosaic-2_{bcom}, and mosaic-2_{ccom}) (**Table 1**). Although RBD4 and RBD7 showed high expression and bound to class 3 and 4, but not class 1 and class 2, anti-RBD antibodies (**Figure 3C**), they were not included because they were designed as sets with RBD3 and RBD8, which had been removed. We also created a mosaic-5_{com} RBD-NP with RBD1, RBD2, RBD4, RBD5, and RBD10 (**Table 1**) to investigate whether immune responses to an RBD-NP containing more RBDs with some overlapping mutations differed from responses to mosaic-2_{com} RBD NPs. RBD6 was excluded because it is similar to RBD10, RBD7 was

excluded because it is similar to RBD1, and RBD9 was excluded because it did not completely abrogate binding of class 1 anti-RBD antibodies (**Figure 3C**). From the zoonotic sarbecovirus RBDs, we used all of the selected RBDs to create a mosaic-7_{COM} RBD-NP, which does not display a SARS-CoV-2 RBD, unlike mosaic-8b RBD-NP (**Table 1**) (27).

Table 1. RBDs in each RBD-NP

Mosaic-2a_{COM}	Mosaic-2b_{COM}	Mosaic-2c_{COM}	Mosaic-5_{COM}	Mosaic-7_{COM}	Mosaic-8b	Mosaic-7	Homotypic SARS-CoV-2
RBD1	RBD5	RBD9	RBD1	LYRa3	Beta	–	Beta
RBD2	RBD6	RBD10	RBD2	Khosta-2	RaTG13	RaTG13	–
–	–	–	RBD4	C028	Pang17	Pang17	–
–	–	–	RBD5	SHC014	SHC014	SHC014	–
–	–	–	RBD10	BM48-31	WIV1	WIV1	–
–	–	–	–	BtKY72	Rs4081	Rs4081	–
–	–	–	–	Pang17	Rf1	Rf1	–
–	–	–	–	–	RmYN02	RmYN02	–

RBDs in computationally designed RBD NPs (mosaic-2_{COMs}, mosaic-5_{COM}, and mosaic-7_{COM}) are defined in Tables S3 and S4. RBDs in mosaic-8b, mosaic-7, and homotypic SARS-CoV-2 are defined in previous studies (27, 43).

Mosaic-8b, mosaic-7, and homotypic SARS-CoV-2 Beta RBD NPs were prepared and characterized as described (26, 27, 43), and conjugations to create mosaic-2a_{COM}, mosaic-2b_{COM}, mosaic-2c_{COM}, mosaic-5_{COM}, and mosaic-7_{COM} were successful, as demonstrated by SEC (**Figure 3E**) and SDS-PAGE (**Figure 3F**). In previous work, we found uniform conjugation efficiencies and near 100% conjugation for all SpyTagged RBDs we evaluated (44).

2.3.4 Designed RBD NPs elicit cross-reactive binding and pseudovirus neutralization responses in naive mice

To assess antibody responses to the designed RBD NPs, we immunized naive BALB/c mice at days 0, 28, and 56 (**Figure 4A**). For mosaic-2_{COM} RBD-NP sequential immunizations, mosaic-2a_{COM} was administered on day 0, mosaic-2b_{COM} on day 28, and mosaic-2c_{COM} on day 56. Mice immunized with three doses of mosaic-8b or homotypic SARS-CoV-2 Beta RBD NPs were included for comparison with other RBD NPs.

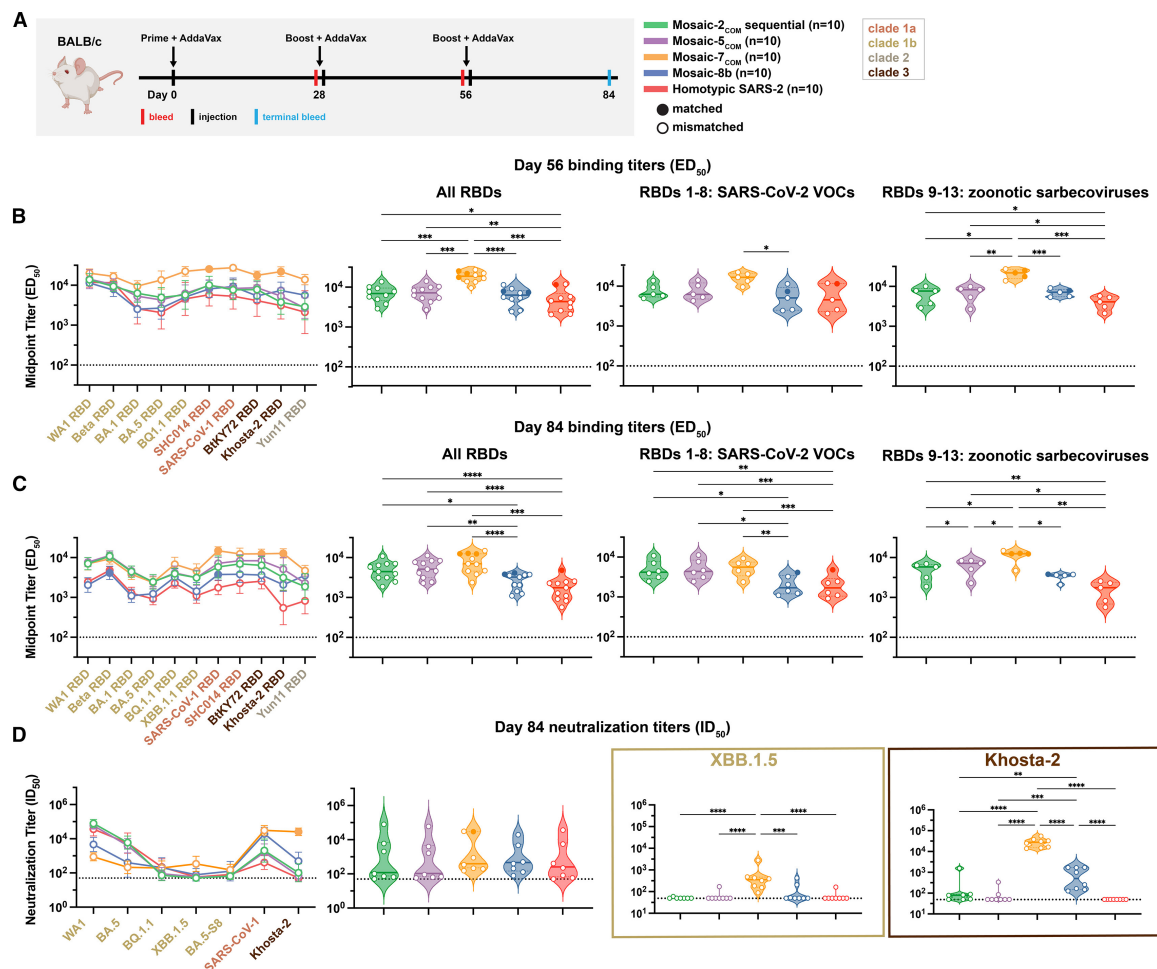


Figure 4. Computationally designed mosaic RBD NPs elicit cross-reactive antibody binding and pseudovirus neutralization responses in immunized mice

The mean of mean titers is compared in (B) and (C) by Tukey's multiple comparison test with the Geisser-Greenhouse correction calculated using GraphPad Prism, with pairings by

viral strain. Error bars associated with each data point are standard deviations. Significant differences between immunized groups linked by horizontal lines are indicated by asterisks: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, $****p < 0.0001$.

(A) Left: schematic of immunization regimen. Middle: numbers and colors used for sarbecovirus strains within clades throughout the figure. Right: colors and symbols (squares) used to identify immunizations (colors) and matched (filled in) versus mismatched (not filled in) viral strains.

(B) Left: ELISA binding titers at day 56 for serum IgG binding to RBDs, represented as mean ED_{50} values. Middle left: means of ELISA binding titers for each immunization. Middle right: means of ELISA binding titers for each immunization against only SARS-CoV-2 variant RBDs. Right: means of ELISA binding titers for each immunization against zoonotic sarbecovirus RBDs. Each circle represents the mean serum IgG binding titer against matched (solid circles) and mismatched (open circles) RBDs.

(C) Left: ELISA binding titers at day 84 for serum IgG binding to RBDs, represented as mean ED_{50} values. Middle left: means of ELISA binding titers for each immunization. Middle right: means of ELISA binding titers for each immunization against only SARS-CoV-2 variant RBDs. Right: means of ELISA binding titers for each immunization against zoonotic sarbecovirus RBDs. Each circle represents the mean serum IgG binding titer against matched (solid circles) and mismatched (open circles) RBDs.

(D) Left: neutralization titers at day 84 for serum IgG neutralization of pseudoviruses derived from the virus strains in (A), represented as mean ID_{50} values. Middle left: means of all neutralization titers for each immunization. Each circle represents the mean neutralization titer against matched (Khosta-2 for mosaic-7_{COM}; solid circle) and mismatched (open circles) pseudoviruses. Middle right and right: neutralization titers against XBB.1.5 and Khosta-2. Each circle represents a neutralization titer from an individual mouse serum sample.

We measured ELISA binding titers against a panel of sarbecovirus RBDs at days 56 and 84 (**Figures 4B and 4C**). Day 56 responses revealed that mosaic-2_{COM} sequential and mosaic-5_{COM} elicited significantly higher titers than homotypic SARS-CoV-2 Beta when comparing means of all RBD titers (**Figure 4B, left**), as well as when comparing mean binding titers

against only zoonotic sarbecovirus strains (**Figure 4B, right**). Interestingly, mosaic-7_{COM} immunization elicited the highest binding titers against all RBDs, including zoonotic sarbecovirus RBDs, rising to significance when comparing mosaic-7_{COM} titers to titers for all other groups.

After three doses of each RBD-NP, the day 84 responses illustrated that the computationally designed RBD NPs consistently elicited significantly higher binding titers against all evaluated RBDs when compared with mosaic-8b and homotypic SARS-CoV-2 (**Figure 4C, left**), an important result because levels of binding antibodies correlate with protection conferred by COVID-19 vaccines (45–47). This was also true when comparing responses against RBDs derived from SARS-CoV-2 variants of concern (VOCs) (**Figure 4C, middle**). However, only the binding responses elicited by mosaic-7_{COM} were significantly better than responses against mosaic-8b when evaluated against zoonotic sarbecovirus RBDs (**Figure 4C, right**).

Although binding antibody responses showed significant differences between cohorts at day 84, mean neutralization titers across evaluated pseudoviruses, all of which were mismatched except for Khosta-2 (matched for mosaic-7_{COM} but not for the other RBD NPs), showed no significant differences (**Figure 4D**). However, mean neutralization titers against individual strains showed some differences (**Figure 4D, left**). For example, mosaic-7_{COM} elicited lower neutralization titers than mosaic-8b against SARS-CoV-2 WA1 and BA.5, likely because mosaic-7_{COM} does not display a SARS-CoV-2 RBD or an RBD that shares >87% sequence identity with the WA1 or BA.5 RBDs (**Figure S3**). However, mosaic-7_{COM} elicited significantly higher neutralization titers against XBB.1.5 (mismatched for all RBD NPs) than the other immunogens (**Figure 4D**) despite lacking a SARS-CoV-2 RBD. As expected, mosaic-7_{COM} also elicited significantly higher neutralization titers against Khosta-2, a matched strain (**Figure 4D**). We also found that mosaic-2_{COM} sequential, mosaic-5_{COM}, and homotypic SARS-CoV-2 Beta RBD NPs elicited higher neutralization titers against SARS-CoV-2 WA1 and BA.5 and lower titers against zoonotic sarbecoviruses than the mosaic-7_{COM} and mosaic-8b RBD NPs, that mosaic-2_{COM} sequential and mosaic-5_{COM} elicited

similar neutralization titers as homotypic SARS-CoV-2 Beta RBD-NP against WA1 and BA.5, and that mosaic-2_{COM} sequential and mosaic-5_{COM} elicited higher or similar titers as homotypic SARS-CoV-2 Beta RBD-NP against non-SARS-CoV-2 sarbecoviruses such as SARS-CoV. However, mosaic-2_{COM} sequential and mosaic-5_{COM} did not elicit detectable neutralization titers against XBB.1.5, suggesting that these RBD NPs may be ineffective against the XBB series and more recent BA.2.86 series. In addition, binding and neutralizing titers elicited by mosaic-2_{COM} sequential and mosaic-5_{COM} immunizations were generally similar to each other.

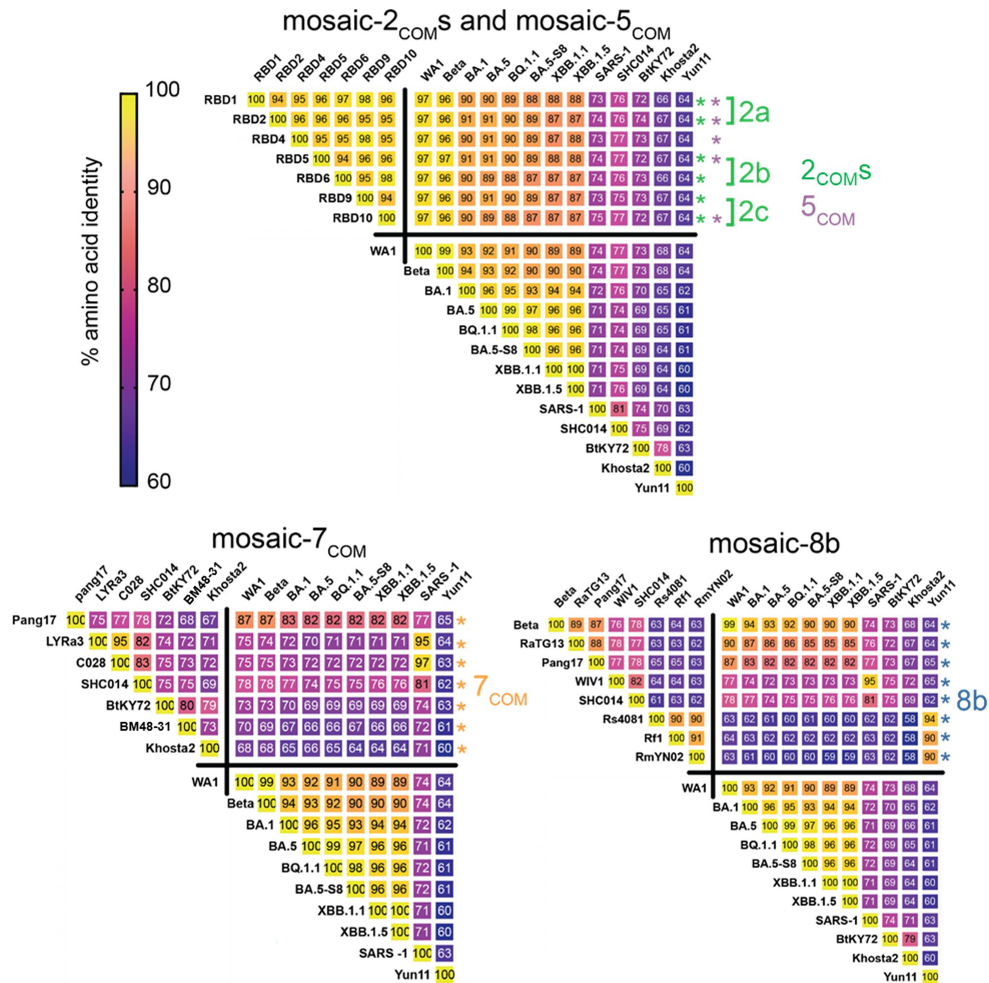


Figure S3. RBD amino acid sequence identities for computationally designed mosaic RBD NPs and mosaic-8b, related to Figure 3

Asterisks indicate strains used for mosaic RBD NPs.

It is interesting that the high binding titers against zoonotic sarbecoviruses elicited by mosaic-2_{COM} and mosaic-5_{COM} were not reflected in their neutralization titers, suggesting that mosaic-2_{COM} and mosaic-5_{COM} elicited non-neutralizing anti-RBD antibodies (e.g., against sterically occluded class 4 RBD epitopes (15)) but fewer class 1/4 anti-RBD antibodies, which tend to be more strongly neutralizing (21).

Taken together, the results suggest that mosaic-2_{COM}s, mosaic-5_{COM}, and especially mosaic-7_{COM} would be effective RBD NPs for eliciting cross-reactive responses in SARS-CoV-2 naive individuals. In addition, these results validate the approach of using mosaic RBD NPs composed of computationally designed or selected zoonotic RBDs to elicit broader antibody binding responses to sarbecoviruses.

2.3.5 DMS reveals targeting of conserved RBD epitopes by mosaic RBD NPs

We further investigated antibody responses raised by mosaic-7_{COM}, which elicited both cross-reactive binding and neutralizing titers against sarbecoviruses from different clades (**Figures 4C and 4D**). To address which RBD epitopes were recognized, we performed DMS using a SARS-CoV-2 Beta yeast display library (19) to compare sera from mice immunized with mosaic-7_{COM}, mosaic-8b, or homotypic SARS-CoV-2 Beta (**Figure 5A**). Consistent with a previous DMS comparison of mosaic-8b and homotypic SARS-CoV-2 Beta DMS profiles (27), we found higher escape values for residues within class 3 and 4 RBD epitopes (epitopes defined as described (15)) and lower escape values in class 2 and class 1 RBD residues for mosaic-7_{COM} and mosaic-8b sera compared with the profile from homotypic SARS-CoV-2 Beta sera. Differences between mosaic-7_{COM} and mosaic-8b sera were difficult to discern across the entire DMS profile but became more apparent when evaluating specific residues on the surface of an RBD (**Figure 5B**). For example, mosaic-7_{COM} showed higher escape than mosaic-8b at residue 383 (a class 4 residue) and residue 360 (a class 3 residue), suggesting that antibodies recognized an epitope involving those sites. Mosaic-7_{COM} serum also showed little to no escape at RBD residue 484, a class 2 residue that showed high escape from both mosaic-8b and homotypic SARS-CoV-2 Beta sera, suggesting that mosaic-7_{COM} elicited fewer class 2 antibodies against SARS-CoV-2 strains.

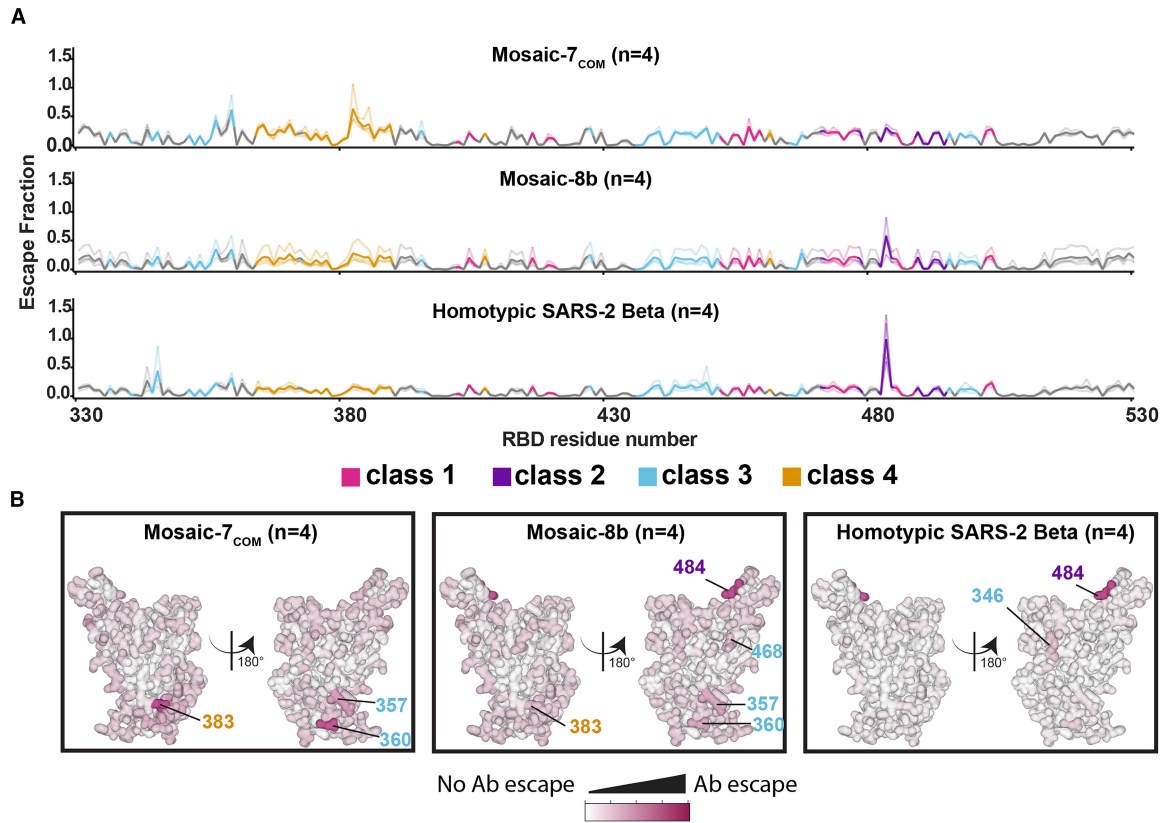


Figure 5. Differences in epitope targeting of antibodies elicited in mice immunized with mosaic and homotypic RBD NPs

(A) DMS line plots for analyses of sera from mice that were immunized as shown in Figure 4A. DMS was conducted using a SARS-CoV-2 Beta RBD library. The x axis shows RBD residue positions, and the y axis shows the total sum of Ab escape for all mutations at a given site, with larger values indicating greater Ab escape. Each faint line represents a single antiserum with heavy lines indicating the average of $n = 4$ sera for each group. Lines are colored differently based on RBD epitopes from the 4 major classes (color definitions are shown in the legend below this panel and gray for residues not assigned to an epitope).

(B) Mean site-total antibody escape for a SARS-CoV-2 Beta RBD library determined using sera from mice immunized with the indicated immunogens mapped to the surface of the WA1 RBD (PDB: 6M0J). White indicates no escape, and dark pink indicates sites with the most escape (residue numbers are denoted with epitope-specific colors as denoted by the legend between A and B).

2.3.6 Mosaic-7_{COM} elicited superior cross-reactive responses in mice with prior COVID-19 vaccinations

We next investigated the impact of prior COVID-19 vaccinations on mosaic-7_{COM} by immunizing BALB/c mice that had previously been vaccinated with two doses of a WA1 Pfizer-equivalent mRNA-LNP vaccine followed by a bivalent WA1/BA.5 mRNA-LNP vaccine (**Figure 6A**). We immunized mice with two doses of mosaic RBD NPs (either mosaic-7_{COM}, mosaic-8b, or mosaic-7, mosaic-8b without SARS-CoV-2 Beta RBD (43)) or an additional dose of bivalent WA1/BA.5 mRNA-LNP. Results for the mosaic-8b and mosaic-7 cohorts in this experiment were previously described (43), and here, we compare those results to mosaic-7_{COM} immunizations because both mosaic-7 RBD NPs lack a SARS-CoV-2 RBD, whereas mosaic-8b includes the SARS-CoV-2 Beta RBD (**Table 1**). As previously discussed, levels of binding antibodies after animals had received the same course of mRNA-LNP vaccines showed significant differences in titers elicited by the pre-vaccinations across cohorts (43) (**Figures S4B and S4C, day 0**). We therefore used baseline corrections (see STAR Methods) to account for different mean responses at day 0 in each of the groups for the data shown in **Figure 6**. (Binding data without baseline corrections are shown in **Figures S4B and S4C**.) Neutralization potencies at day 0 were similar for all cohorts and therefore were not baseline corrected (43).

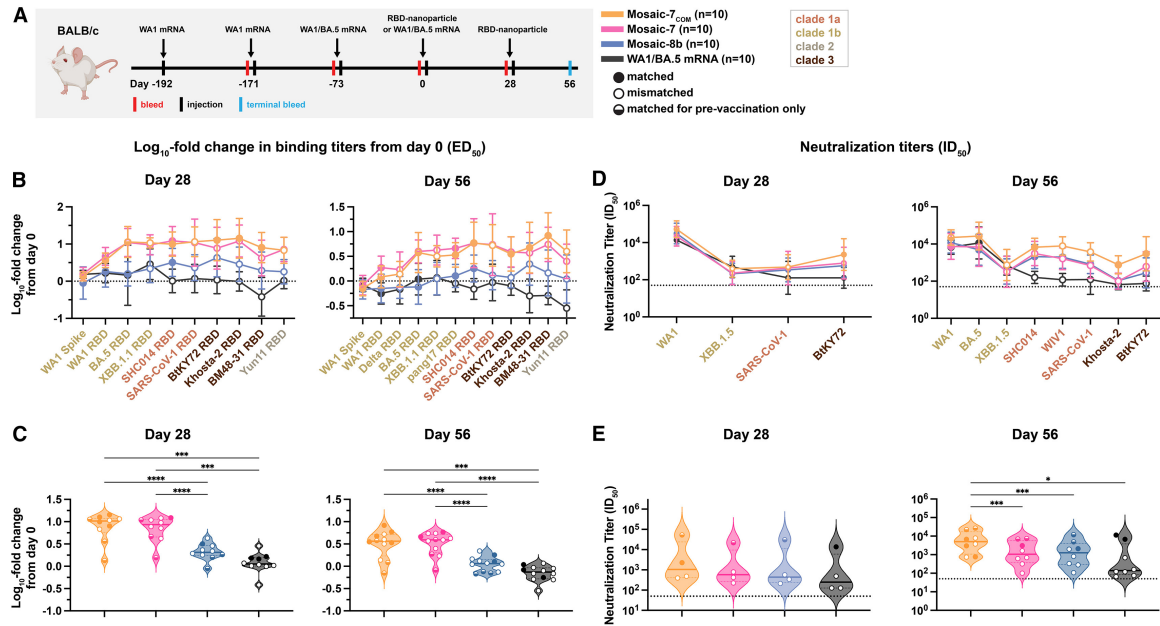


Figure 6. Mosaic-7_{COM} immunization in pre-vaccinated mice elicits superior cross-reactive antibody responses

The mean of mean titers is compared in (C) and (E) by Tukey's multiple comparison test with the Geisser-Greenhouse correction calculated using GraphPad Prism, with pairings by viral strain. Error bars associated with each data point are standard deviations. Significant differences between immunized groups linked by horizontal lines are indicated by asterisks: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$. Binding responses at day 0 (before NP or other vaccine immunizations) showed significant differences across cohorts in titers elicited by the pre-vaccinations (43). To account for different mean responses at day 0 between cohorts, we applied baseline corrections (see STAR Methods). Uncorrected binding data for (B) and (C) are shown in Figures S4B and S4C.

(A) Left: schematic of vaccination regimen. Mice were pre-vaccinated with mRNA-LNP encoding WA1 spike and bivalent WA1/BA.5 prior to prime and boost immunizations with RBD NPs at day 0 and day 28 or an additional WA1/BA.5 mRNA-LNP immunization at day 0. Middle: colors and symbols (squares) used to identify immunizations (colors) and matched (filled in), mismatched (not filled in), or matched to pre-vaccination (half-filled in) viral strains (squares). Right: numbers and colors used for sarbecovirus strains within clades throughout the figure.

(B) Log₁₀ mean fold change in ELISA ED₅₀ binding titers from day 0 at the indicated days after priming with the indicated immunogens against spike or RBD proteins from the indicated sarbecovirus strains (numbers and color coding as in A).

(C) Log₁₀ means of fold change in ELISA titers for each type of immunization at the indicated days. Each circle represents the log₁₀ mean fold change in ED₅₀ titers from mice against a single viral strain of sera from mice that were immunized with the specified immunogen (solid circles = matched; open circles = mismatched; colors for different strains defined in A).

(D) Mean change in neutralization ID₅₀ titers from day 0 at the indicated days against the indicated sarbecovirus strains (numbers and color coding as in A).

(E) Means of all neutralization titers for each type of immunization at the indicated days. Each circle represents the mean neutralization IC₅₀ titer against a single viral strain of sera from mice that were immunized with the specified immunogen (solid circles = matched; open circles = mismatched; colors for different strains defined in A).

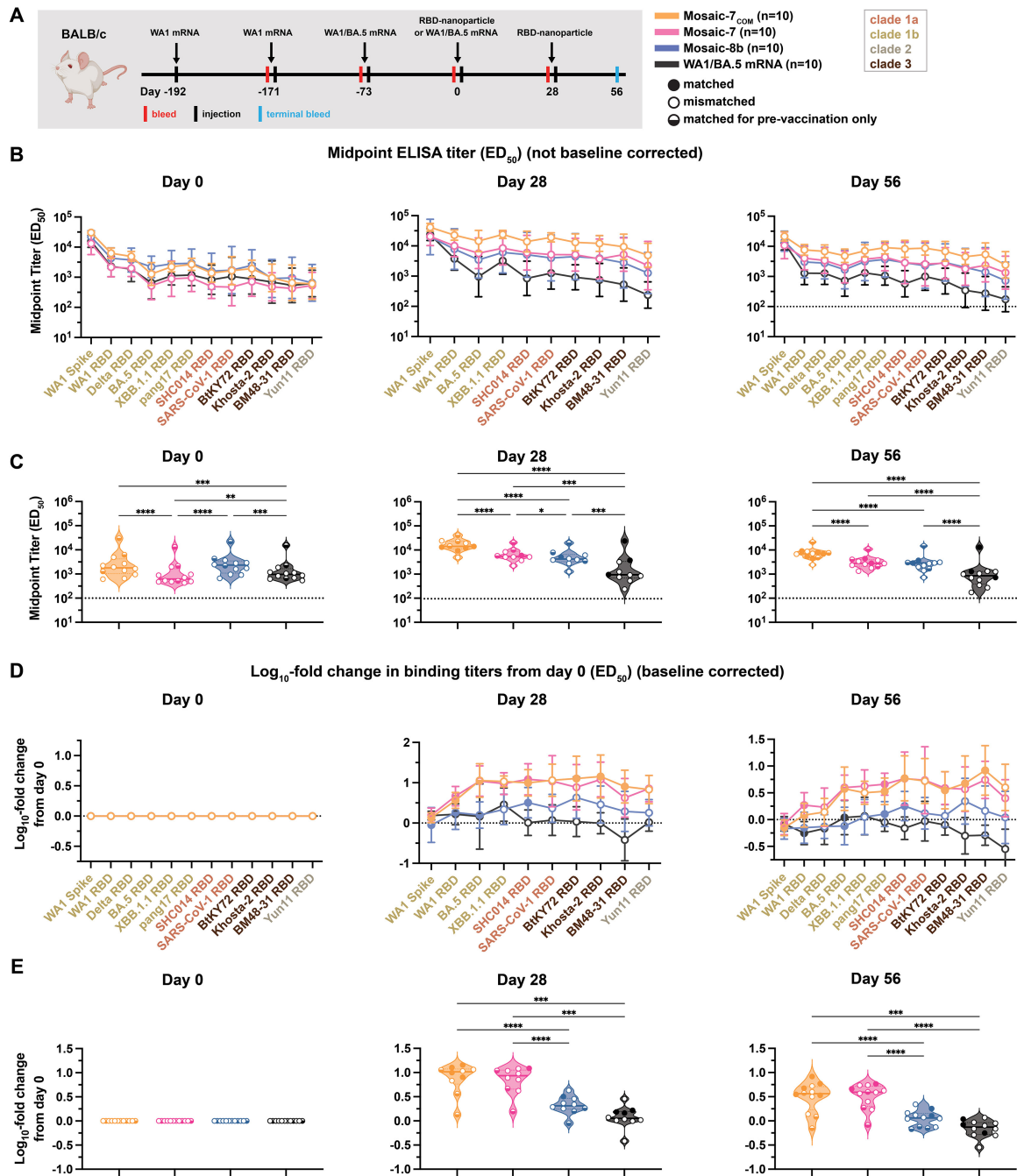


Figure S4. Mosaic-7_{COM} immunization in pre-vaccinated mice elicited superior cross-reactive antibody responses, related to Figure 6

The mean of mean titers is compared in (C) and (E) by Tukey's multiple comparison test with the Geisser-Greenhouse correction calculated using GraphPad Prism, with pairings by viral strain. Error bars associated with each data point are standard deviations. Significant

differences between immunized groups linked by horizontal lines are indicated by asterisks: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, $****p < 0.0001$. Binding responses at day 0 (before NP or other vaccine immunizations) showed significant differences across cohorts in titers elicited by the pre-vaccinations (43). To account for different mean responses at day 0 between cohorts, we applied baseline corrections in Figure 6 (see STAR Methods). Here, binding data are shown as both baseline corrected (B and C) and not baseline corrected (D and E).

(A) Left: schematic of vaccination regimen. Mice were pre-vaccinated with mRNA-LNP encoding WA1 spike and bivalent WA1/BA.5 prior to prime and boost immunizations with RBD NPs at day 0 and day 28 or an additional WA1/BA.5 mRNA-LNP immunization at day 0. Middle: colors and symbols (squares) used to identify immunizations (colors) and matched (filled in), mismatched (not filled in), or matched to pre-vaccination (half-filled in) viral strains (squares). Right: numbers and colors used for sarbecovirus strains within clades throughout the figure.

(B) ELISA ED_{50} binding titers in serum samples from mice immunized with the indicated immunogens measured at days 0, 28, and 56 against spike or RBD proteins from the indicated sarbecovirus strains (numbers and color coding as in A).

(C) Mean ELISA titers for each type of immunization at the indicated days. Each circle represents the mean ED_{50} titers from mice against a single viral strain of sera from mice that were immunized with the specified immunogen (solid circles = matched; open circles = mismatched; colors for different strains defined in A).

(D) Mean fold change in ELISA ED_{50} binding titers from day 0 in serum samples from mice immunized with the indicated immunogens measured at days 0, 28, and 56 against spike or RBD proteins from the indicated sarbecovirus strains (numbers and color coding as in A).

(E) Means of fold changes in ELISA titers for each type of immunization at the indicated days. Each circle represents the mean fold change in ED_{50} titers from mice against a single viral strain of sera from mice that were immunized with the specified immunogen (solid circles = matched; open circles = mismatched; colors for different strains defined in A).

Days 28 and 56 log₁₀-fold changes in ELISA binding titers (after prime and boosting RBD-NP immunizations) are shown in **Figure 6B**. At both time points, RBD NPs generally more strongly boosted binding titers than a second dose of a bivalent WA1/BA.5 mRNA-LNP vaccine, especially against zoonotic sarbecoviruses. Mosaic-8b boosted titers less than mosaic-7 and mosaic-7_{COM} against all viral strains, and mosaic-7 and mosaic-7_{COM} largely boosted titers to similar extents. Mean log₁₀-fold changes in binding titers (**Figure 6C**) showed that mosaic-7 and mosaic-7_{COM} both boosted binding titers significantly better than mosaic-8b or WA1/BA.5 mRNA-LNP.

Mosaic-7 and mosaic-8b elicited similar neutralization titers at both days 28 and 56 (**Figure 6D**), whereas mosaic-7_{COM} elicited higher neutralization titers than other immunogens against both zoonotic sarbecoviruses and SARS-CoV-2 variants at day 56. Notably, mosaic-7_{COM} antisera neutralized XBB.1.5 with equal potency as antisera from mice boosted with a second dose of WA1/BA.5 mRNA-LNP, whereas antisera from mosaic-7 and mosaic-8b prime/boosted animals exhibited lower potencies, resulting in statistically significant differences in mean neutralization titers between mosaic-7_{COM} and other immunogens (**Figure 6E**).

Although the mean of means of baseline-corrected binding titers for mosaic-7_{COM} and mosaic-7 cohorts were similar to each other (**Figure 6B**) and both were significantly higher than titers for mosaic-8b (**Figure 6C**), the non-baseline-corrected-binding titers for these groups at days 28 and 56 showed differences (**Figure S4B**): e.g., mosaic-7_{COM} sera elicited significantly higher mean of mean titers than mosaic-7 or mosaic-8b (**Figure S4C**). The higher titers for mosaic-7_{COM} compared with mosaic-7 could be related to the fact that the mosaic-7_{COM} cohort started with higher antibody binding titers than mosaic-7 at day 0 (prior to RBD-NP immunizations), but this does not apply to differences with the mosaic-8b cohort since the mosaic-7_{COM} and mosaic-8b titers at day 0 were equivalent (**Figure S4C**).

Taking all experimental results into account, mosaic-7_{COM} showed broader and more potent binding and neutralization than either of its mosaic NP counterparts, suggesting that mosaic-

7_{COM} more efficiently elicits broader binding and more potently neutralizing antibodies in a pre-vaccinated animal model.

2.4 Discussion

Multivalent NPs have emerged as a useful platform for developing new vaccines against mutable pathogens, including influenza (44, 48), human immunodeficiency virus (HIV) (49–51), respiratory syncytial virus (RSV) (50), and SARS-CoV-2 (26, 27, 43, 52–60). Many of these are homotypic SARS-CoV-2 spike- or RBD NPs, which display multiple copies of only one SARS-CoV-2 spike or RBD and are analogous to homotypic SARS-CoV-2 Beta RBD-NP, which we used for previous and current comparisons with mosaic-8b RBD NPs (27, 43). Displaying different variants on a single NP could provide broader protection, as previous studies demonstrated that the mosaic-8b RBD-NP protected K18-hACE2 transgenic mice from a mismatched SARS-CoV challenge, whereas homotypic SARS-CoV-2 Beta RBD-NP did not (27). Recently, we also studied the impact of prior COVID-19 vaccinations on mosaic-8b RBD-NP vaccines, finding that mosaic-8b still elicited cross-reactive antibodies, primarily by boosting them from prior vaccinations (43).

The “plug-and-display” SpyCatcher-SpyTag system (24, 25) has been used with a variety of antigens (26, 27, 44, 61–63) and can be easily adapted to make mosaic NPs with different antigenic compositions. This flexibility led us to evaluate whether designing RBD sequences using computational methods and available data could enhance elicited cross-reactive antibody responses beyond those of previously studied RBD NPs such as mosaic-8b, mosaic-7, or homotypic SARS-CoV-2 Beta (43). We designed two sets of RBD NPs that are effective for different targets. First, we combined DMS data, machine learning, and structure-based solubility predictions to design RBDs with 6 mutations relative to the WA1 RBD, which were then displayed as sets of 2 RBDs on NPs. RBDs within a set were designed to reduce bivalent B cell receptor binding to class 1 and 2 RBD epitopes while maintaining bivalent binding to class 3 and 4 epitopes, and these designed RBDs were then used to create the mosaic-2_{COMS} and mosaic-5_{COM} RBD NPs. Since these RBDs only included a few substitutions relative to the WA1 RBD, the mosaic-2_{COMS} and mosaic-5_{COM} were designed

to be effective against less mutated SARS-CoV-2 variants. Second, we used DMS escape measurements (17, 30–33) and sequence diversity to select naturally occurring sarbecovirus RBDs to create the mosaic-7_{COM} RBD-NP, which was designed to be effective against zoonotic sarbecoviruses and heavily mutated SARS-CoV-2 variants. Binding and neutralization titers following immunization of naive mice suggested that the designed RBD NPs are indeed most effective against their respective proposed targets, and all of the computationally designed RBD NPs were superior to previously described RBD NPs (homotypic SARS-CoV-2 Beta for less mutated SARS-CoV-2 variants and mosaic-8b/mosaic-7 for zoonotic sarbecoviruses and heavily mutated SARS-CoV-2 variants) (43).

Although less mutated SARS-CoV-2 variants are no longer circulating in humans, our computational method for generating mosaic-2_{COMS} and mosaic-5_{COM} could still be useful. For example, the mosaic-2_{COMS} and mosaic-5_{COM} were just as potent as homotypic SARS-CoV-2 Beta RBD-NP against less mutated SARS-CoV-2 variants and more potent against zoonotic sarbecoviruses and heavily mutated SARS-CoV-2 VOCs. Thus, the mosaic-2_{COMS} and mosaic-5_{COM} exhibited superior properties compared with homotypic Beta RBD-NP. Our designs had an 80% success rate of producing folded SpyTagged RBDs that expressed well, suggesting that these methods could be used for other SARS-CoV-2 variants or for other viruses.

To address the ongoing rise of SARS-CoV-2 Omicron variants and potential zoonotic sarbecovirus spillovers, we introduce mosaic-7_{COM} as an RBD-NP that provides more effective cross-reactive responses than other mosaic RBD NPs (mosaic-8b and mosaic-7) (27, 43) in both naive and pre-vaccinated mice. DMS results, although possibly obscured by the presence of multiple antibody classes in polyclonal sera (43), suggested that mosaic-7_{COM} elicited more antibodies binding to class 3 and 4 epitopes than mosaic-8b and homotypic SARS-CoV-2 Beta RBD NPs. Additionally, mosaic-7_{COM} elicited fewer antibodies recognizing an epitope involving the class 2 RBD residue 484, suggesting that mosaic-7_{COM} effectively redirects the antibody response from variable epitopes to conserved epitopes compared with mosaic-8b. We note that both of the mosaic RBD NPs that lack a SARS-

CoV-2 RBD (mosaic-7_{COM} and mosaic-7) outperformed mosaic-8b in mice that had received mRNA-LNP vaccines expressing SARS-CoV-2 WA1 and Omicron BA.5 spikes, consistent with a lack of expansion of SARS-CoV-2-specific immunodominant antibodies that target variable epitopes in mice receiving RBD NPs that did not include a SARS-CoV-2 RBD. Additionally, mosaic-7_{COM} and mosaic-7 outperform mosaic-8b despite the possibility that pre-existing high-affinity antibodies could block variable SARS-CoV-2 epitopes and enhance cross-reactive responses to mosaic-8b (34, 64–68). Although removal of a SARS-CoV-2 RBD from mosaic-8b to create mosaic-7 improved binding responses in pre-immunized mice, mosaic-7_{COM} elicited significantly higher neutralizing titers than mosaic-7 against both zoonotic sarbecoviruses and SARS-CoV-2 variants, supporting its use as a pan-sarbecovirus vaccine in populations that have been exposed to SARS-CoV-2.

Extending the mosaic NP approach to other mutable viruses could be an interesting avenue of research, as promoting avidity advantages of cross-reactive B-cell receptors (BCRs) to enhance cross-reactive responses may not be unique to sarbecoviruses. A monomeric form of viral antigen may be necessary in some cases to avoid identical epitopes being displayed on viral trimers that could be bound bivalently by strain-specific BCRs (44, 69), thereby eliciting a higher proportion of antibodies against variable, rather than conserved, epitopes. In addition, modeling of antigen-NPs with IgG models derived from known structures of fragment antigen-binding (Fab) regions from mAbs against conserved regions is advisable to ensure that conserved epitopes are accessible on antigen-NPs, e.g., as shown for models of mAb IgGs elicited against mosaic RBD NPs (70).

A guiding principle for creating effective mosaic RBD NPs is maximizing the diversity of RBDs displayed on a single NP to focus the response on conserved epitopes (27, 69). This can be done by increasing the number of variant RBDs to decrease the probability that B cell receptors can crosslink between immunodominant variable epitopes on adjacent identical RBDs. In this work, the enhanced cross-reactive responses elicited by mosaic-7_{COM} illustrate that computationally optimizing protein sequences is another way to increase displayed RBD diversity. Overall, our results support the integration of computational methods with vaccine

design and the further evaluation of designed RBD NPs, particularly mosaic-7_{COM}, as potential pan-sarbecovirus vaccines.

2.4.1 Limitations of the study

Our computational designs for the RBDs in mosaic-2_{COMS} and mosaic-5_{COM} were limited to 6 mutations due to the limited number of mutations found in the DMS training data for our expression predictor (29). Although we have shown that the RBD NPs reported here remain stable long enough to elicit cross-reactive responses, we have not systematically examined long-term stability of these RBD NPs as would be required for approval of a vaccine candidate for human use.

The SARS-CoV-2 VOCs included in our binding and pseudovirus neutralization assays represented the most recent VOCs at the time of the experiments, but we could not assess more recent VOCs due to limited amounts of antiserum. We also could not perform authentic virus neutralization assays due to a lack of appropriate facilities and limited antiserum, but this issue is mitigated by our previous study that found quantitative consistency between pseudovirus neutralization titers and authentic virus neutralization titers from mosaic RBD-NP immunization (27) and by an earlier comprehensive comparison of SARS-CoV-2 pseudovirus-based and authentic virus neutralization that reached a similar conclusion (71). Due to the unavailability of licensed Pfizer-BioNTech or Moderna mRNA-LNP vaccines for research, we utilized mRNA-LNP formulations from alternative sources, which may not perform identically to the clinically approved vaccines.

In our DMS assay, the SARS-CoV-2 Beta RBD library was mismatched for mosaic-7_{COM} but matched for mosaic-8b and homotypic SARS-CoV-2 Beta, so there was a greater chance of observing signals in conserved epitopes for mosaic-7_{COM}. A matched comparison for both mosaic-7_{COM} and homotypic SARS-CoV-2 Beta could not be made since mosaic-7_{COM} does not display a SARS-CoV-2 Beta RBD. In addition, we previously observed that polyclonal antisera containing antibodies of multiple RBD classes (“polyclass” antibodies) tend to have obscured DMS signals and low escape fractions over all residues compared with DMS

signals from mAbs or mixtures of anti-RBD antibodies in which one antibody class is dominant (43). Thus, it is possible that the differences between DMS profiles for mosaic-7_{COM} and mosaic-8b were dampened by the polyclonal nature of the elicited antibodies against the mosaic RBD NPs.

2.5 Materials and Methods

2.5.1 Mice

6- to 7-week-old female BALB/c mice (Charles River Laboratories) were housed at Labcorp Drug Development, Denver, PA for immunizations. All animals were healthy after being weighed and monitored for 7 days preceding the start of the study. Mice were randomly assigned to experimental groups of 10. Cages were kept in a climate-controlled room at 68-79°C at 50 ± 20% relative humidity. Mice were provided Rodent Diet #5001 (Purina Lab Diet) ad libitum. Mouse procedures were approved by the Labcorp Institutional Animal Care and Use Committee.

2.5.2 Protein expression and purification

Monoclonal IgGs, a soluble SARS-CoV-2 trimer with 6P stabilizing mutations, and a human ACE2-Fc construct (72) were produced as previously described (21, 26, 27).

Vectors encoding the protein sequences for computationally designed RBDs were assembled using Gibson cloning from an insert encoding residues 319–541 of the SARS-CoV-2 WA1 RBD with the indicated substitutions (Tables 1 and S3). For the selected sarbecovirus RBDs in Table S4, RBDs were also assembled using Gibson cloning from inserts encoding the indicated residues. RBDs used for mosaic-8b, mosaic-7, and homotypic SARS-2 were expressed as described previously (43). Sarbecovirus RBDs from SARS-CoV-2 Beta (GenBank: QUT64557.1), SARS-CoV-2 WA1 (GenBank: MN985325.1), SARS-CoV-2 BA.5-S8 (73), LYRa3 (GenBank: AHX37569.1) (73), Khosta-2 CoV (GenBank: QVN46569.1), SHC014-CoV (GenBank: KC881005), BM48-31-CoV (GenBank: NC014470), BtKY72-CoV (GenBank: KY352407), Yun11-CoV (GenBank: JX993988), WIV1-CoV (GenBank: KF367457), RaTG13-CoV (GenBank: QHR63300), SARS-CoV

(GenBank: AAP13441.1), Rs4081-CoV (GenBank: KY417143), RmYN02-CoV (GSAID: EPI_ISL_412977), Rf1-CoV (GenBank: DQ412042), and pangolin17-CoV (GenBank: QIA48632) were constructed as previously described (15, 26, 27, 43). Briefly, RBDs used for conjugations were encoded with a C-terminal hexahistidine tag (6xHis; G-HHHHHH) and SpyTag003 (RGVPHIVMVDAYKRYK) (74) for conjugating onto SpyCatcher003-mi3 to form mosaic NPs. RBDs used for ELISAs were encoded with a C-terminal Avi tag (GLNDIFEAQKIEWHE) followed by a hexahistidine tag (6xHis; G-HHHHHH).

RBDs were expressed and subsequently purified via His-tag affinity and SEC purification from transiently-transfected Expi293F (ThermoScientific) supernatants (27). RBDs for immunizations in mRNA-LNP pre-vaccinated mice were prepared as described for the analogous experiment (43).

2.5.3 Preparation of RBD-NPs

SpyCatcher003-mi3 nanoparticles for the RBD-NPs were expressed and purified as described (43). Equivalent conjugation of each RBD in mosaic RBD-NPs was verified by SEC and SDS-PAGE analysis of conjugations to make homotypic NPs for each RBD as described (26, 27, 43). NPs were aliquoted and flash frozen in liquid nitrogen before being stored at -80°C until use.

For immunizations in naïve mice, SpyTagged RBDs were conjugated onto SpyCatcher003-mi3 as described (26, 27). Briefly, equimolar amounts of RBDs for each mosaic NP were mixed before addition of purified SpyCatcher003-mi3, with the final concentration having 2-fold molar excess of total RBD to mi3 subunit: an equimolar mixture of 2 RBDs for the mosaic-2_{COMS}, 5 RBDs for mosaic-5_{COM}, 7 RBDs for mosaic-7_{COM} and mosaic-7, 8 RBDs for mosaic-8b, or only SARS-CoV-2 Beta for homotypic RBD-NPs. Reactions were incubated overnight at room temperature in Tris-buffered saline (TBS) on an orbital shaker. Free RBDs were purified the next day by SEC on a Superose 6 10/300 column (GE Healthcare) and equilibrated with PBS (20 mM sodium phosphate pH 7.5, 150 mM NaCl). RBD-NP conjugations were assessed by SDS-PAGE. Concentrations of conjugated mi3

nanoparticles are reported based on RBD content, determined using a Bio-Rad Protein Assay. RBD-NPs were aliquoted and flash frozen in liquid nitrogen before being stored at -80°C until use.

RBD-NPs for immunizations in mRNA-LNP pre-vaccinated mice were prepared as described for the analogous experiment (43).

2.5.4 Immunizations

For immunizations in naïve animals, RBD-NPs were diluted using Dulbecco's PBS and mixed 1:1 (v/v) with AddaVax prior to immunization, for a final vaccine dose of 5 ug of RBD equivalents in 0.1 mL total volume. RBD-NP immunizations were administered intramuscularly (IM) via both right and left hindleg (50 µL each). Mice were immunized three times at days 0, 28, and 56, bled via tail vein at days 0, 28, and 56, with a terminal bleed via cardiac puncture at day 84. Blood samples were allowed to clot, and sera were collected and stored at -80°C prior to use.

As previously described (43), mice used in the pre-vaccination study were vaccinated IM with 20 µL of WA1 mRNA-LNP at days -192 and -171 containing 1 µg mRNA diluted in PBS and 20 µL of WA1/BA.5 mRNA-LNP (0.5 µg WA1 and 0.5 µg BA.5 mRNA) at day -73. Mice were then immunized IM with 5 µg of protein nanoparticle (RBD equivalents) in 100 µL containing 50% v/v AddaVax adjuvant on days 0 and 28 or received an additional dose of 1 µg WA1/BA.5 mRNA-LNP at day 0. Mice were bled and sera were collected as described above.

2.5.5 Reagents used for pre-vaccinations

Pfizer-equivalent mRNA-LNP formulations for WA1 and BA.5 were purchased from Helix Biotech as described (43). Bivalent WA1/BA.5 mRNA LNP was prepared by mixing WA1 mRNA-LNP and BA.5 mRNA-LNP 1:1 by mRNA mass.

2.5.6 Antibody binding and pseudovirus neutralization assays

Binding of characterized anti-RBD monoclonal antibodies and a human ACE2-Fc construct to RBDs was assessed as described (21). Monoclonal antibodies were assigned to RBD epitopes based on structural studies as described (15).

Binding to purified RBDs or spike proteins was assessed using serum samples from immunized mice by ELISA as described (43). We used Graphpad Prism 10.1.1 to plot and analyze binding curves, assuming a one-site binding model with a Hill coefficient to obtain midpoint titers (ED_{50} values for serum ELISAs, EC_{50} values for monoclonal antibody ELISAs). ED_{50}/EC_{50} values were normalized and mean ED_{50}/EC_{50} values were calculated as described (43). For pre-vaccinated mouse data shown in **Figure 6**, ED_{50} s were normalized by dividing the ED_{50} response at day 28 and day 56 (after RBD-NP immunizations or an additional mRNA-LNP immunization) over the ED_{50} response at day 0 to account for differences in binding responses between groups in the pre-vaccination cohorts (**Figure S4C, left**). **Figure S4** compares non-baseline correcting binding responses across cohorts (panels B and C) with baseline corrected binding responses (panels D and E).

Lentiviral-based pseudoviruses were prepared and neutralization assays conducted and luciferase activity was measured as relative luminescence units (RLUs) as described (43). In the context of SARS-CoV-2, pseudovirus neutralization titers have been found to quantitatively correlate with authentic virus neutralization titers in antisera from RBD-NP immunizations (27), convalescent plasma (71), and monoclonal antibodies (71). Relative RLUs were normalized to RLUs from cells infected with pseudotyped virus in the absence of antiserum. Half-maximal inhibitory dilutions (ID_{50} values) were derived using 4-parameter nonlinear regression in Antibody Database (75).

2.5.7 DMS

DMS studies used to map epitopes recognized by serum Abs were performed in biological duplicates using independent SARS-CoV-2 Beta-based mutant libraries (generously provided by Tyler Starr, University of Utah) as described previously (19). Serum samples

were heat inactivated and depleted for yeast binding as described before (43). DMS was performed and escape fractions were calculated and analyzed as described (43, 76). Raw sequencing data are available on the NCBI SRA under BioProject: PRJNA1067836 and BioSample: SAMN46235696.

Static line plot visualizations of escape maps were created using Swift DMS as described (43, 76). Line heights in static line plot visualizations of escape maps (created as described (43, 76)) indicate the escape score for that amino acid mutation. RBD epitopes were classified using previously-described class 1, 2, 3, and 4 nomenclature (15). For structural visualizations, an RBD surface of PDB: 6M0J was colored by the site-wise escape metric at each escape site, with dark pink scaled to be the maximum escape fraction used to scale the y-axis for serum Abs and white indicating no escape. Residues that exhibited the greatest escape fractions were marked with their residue number and colored according to RBD epitope class.

2.5.8 Quantification and statistical analysis

We evaluated whether differences in overall binding titers ($\log_{10}$ ED₅₀s) or pseudovirus neutralization titers ($\log_{10}$ ID₅₀s) between immunized groups of mice (**Figures 4B–4D, 6C, 6E, S4C, and S4E**) were statistically significant using analysis of variance (ANOVA) followed by Tukey's multiple comparison post hoc tests with the Geisser-Greenhouse correction in GraphPad Prism 10.1.0. We evaluated whether differences in pseudovirus neutralization titers between different immunized groups of mice for individual strains were statistically significant (**Figure 4D, XBB.1.5 and Khosta-2**) using analysis of variance (ANOVA) followed by Tukey's multiple comparison with a single pooled variance.

2.6 Resource Availability

2.6.1 Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Arup K. Chakraborty (arupc@mit.edu), who will work with the collaborating Bjorkman lab when necessary.

2.6.2 Materials availability

All unique/stable reagents generated in this study will be made available on request by the lead contact with a completed materials transfer agreement (who will work with the collaborating Bjorkman lab when necessary).

2.6.3 Data and code availability

All original code and data files for the computational designs have been deposited at https://github.com/ericzwang/designed_mosaic_NPs and are publicly available. A Swift DMS program for processing and visualizing DMS data is available from authors upon request.

2.7 Acknowledgments

This work was supported by the National Science Foundation Graduate Research Fellowship: 1745302 (E.W.), the National Institutes of Health: 1-R61-AI161805 and U19AI057229 (A.K.C.), the National Institutes of Health: P01-AI165075 (P.J.B.), Wellcome Leap (P.J.B.), the Bill and Melinda Gates Foundation: INV-034638 (P.J.B.), the Coalition for Epidemic Preparedness Innovations (CEPI) (P.J.B.), and the Merkin Institute for Translational Research (Caltech). E.W. and A.K.C. acknowledge the MIT SuperCloud and Lincoln Laboratory Supercomputing Center for providing HPC resources that have contributed to the research results reported within this work. We thank Jesse Bloom (Fred Hutchinson), Allie Greaney (University of Washington), and Tyler Starr (University of Utah) for RBD libraries and help setting up DMS at Caltech; Justin Bois (Caltech) for advice about statistical analyses; Jost Vielmetter, Luisa Segovia, Annie Lam, and the Caltech Beckman Institute Protein Expression Center for protein production; Igor Antoshechkin and the Caltech Millard and Muriel Jacobs Genetics and Genomics Laboratory for Illumina sequencing; Chengcheng Fan for generating RBD-NP models; Labcorp Drug Development (Denver, PA) for performing mouse immunizations; and Anthony West for calculating the probabilities of identical neighboring RBDs.

2.8 Author Contributions

Conceptualization, E.W., A.K.C., A.A.C., and P.J.B.; methodology, E.W. and A.K.C. (computations); J.R.K., A.V.R., Y.M.A., and P.N.P.G. (experiments); computation and software, E.W.; investigation, E.W. (computations); J.R.K., A.V.R., Y.M.A., and P.N.P.G. (experiments); writing – original draft, E.W., A.K.C., A.A.C., L.F.C., and P.J.B.; writing – review and editing, E.W., A.K.C., A.A.C., L.F.C., and P.J.B.; visualization, E.W., A.A.C., and L.F.C.; supervision, A.K.C. and P.J.B.; project administration, A.K.C. and P.J.B.; funding, A.K.C. and P.J.B.

2.9 Declaration of Interests

A.K.C. is a consultant (titled “Academic Partner”) for Flagship Pioneering, a consultant and member of the Board of Directors of its affiliated company, Apriori Bio, and is a consultant and Scientific Advisory Board Member of another affiliated company, Metaphore Bio. He holds equity in these companies and Dewpoint Therapeutics. P.J.B. and A.A.C. are inventors on a US patent application (17/523,813) filed by the California Institute of Technology that covers mosaic RBD NPs. P.J.B. is a scientific advisor for Vaccine Company, Inc. and for the Vir Biotechnology.

2.10 References

1. Y. Cao, A. Yisimayi, F. Jian, W. Song, T. Xiao, L. Wang, S. Du, J. Wang, Q. Li, X. Chen, Y. Yu, P. Wang, Z. Zhang, P. Liu, R. An, X. Hao, Y. Wang, J. Wang, R. Feng, H. Sun, L. Zhao, W. Zhang, D. Zhao, J. Zheng, L. Yu, C. Li, N. Zhang, R. Wang, X. Niu, S. Yang, X. Song, Y. Chai, Y. Hu, Y. Shi, L. Zheng, Z. Li, Q. Gu, F. Shao, W. Huang, R. Jin, Z. Shen, Y. Wang, X. Wang, J. Xiao, X. S. Xie, BA.2.12.1, BA.4 and BA.5 escape antibodies elicited by Omicron infection. *Nature* **608**, 593–602 (2022).
2. N. P. Hachmann, J. Miller, A. Y. Collier, J. D. Ventura, J. Yu, M. Rowe, E. A. Bondzie, O. Powers, N. Surve, K. Hall, D. H. Barouch, Neutralization Escape by SARS-CoV-2 Omicron Subvariants BA.2.12.1, BA.4, and BA.5. *N. Engl. J. Med.* **387**, 86–88 (2022).
3. T. A. Bates, H. C. Leier, Z. L. Lyski, S. K. McBride, F. J. Coulter, J. B. Weinstein, J. R. Goodman, Z. Lu, S. A. R. Siegel, P. Sullivan, M. Strnad, A. E. Brunton, D. X. Lee, A. C. Adey, B. N. Bimber, B. J. O’Roak, M. E. Curlin, W. B. Messer, F. G. Tafesse, Neutralization of SARS-CoV-2 variants by convalescent and BNT162b2 vaccinated serum. *Nat. Commun.* **12**, 5135 (2021).
4. D. Planas, D. Veyer, A. Baidaliuk, I. Staropoli, F. Guivel-Benhassine, M. M. Rajah, C. Planchais, F. Porrot, N. Robillard, J. Puech, M. Prot, F. Gallais, P. Gantner, A. Velay, J. Le Guen, N. Kassis-Chikhani, D. Edriss, L. Belec, A. Seve, L. Courtellemont, H. Péré, L. Hocqueloux, S. Fafi-Kremer, T. Prazuck, H. Mouquet, T. Bruel, E. Simon-Lorière, F. A. Rey, O. Schwartz, Reduced sensitivity of SARS-CoV-2 variant Delta to antibody neutralization. *Nature* **596**, 276–280 (2021).
5. D. Zhou, W. Dejnirattisai, P. Supasa, C. Liu, A. J. Mentzer, H. M. Ginn, Y. Zhao, H. M. E. Duyvesteyn, A. Tuekprakhon, R. Nutalai, B. Wang, G. C. Paesen, C. Lopez-Camacho, J. Slon-Campos, B. Hallis, N. Coombes, K. Bewley, S. Charlton, T. S. Walter, D. Skelly, S. F. Lumley, C. Dold, R. Levin, T. Dong, A. J. Pollard, J. C. Knight, D. Crook, T. Lambe, E. Clutterbuck, S. Bibi, A. Flaxman, M. Bittaye, S. Belij-Rammerstorfer, S. Gilbert, W. James, M. W. Carroll, P. Klenerman, E. Barnes, S. J. Dunachie, E. E. Fry, J. Mongkolsapaya, J. Ren, D. I. Stuart, G. R. Screaton, Evidence of escape of SARS-CoV-2 variant B.1.351 from natural and vaccine-induced sera. *Cell* **184**, 2348–2361.e6 (2021).
6. B. J. Willett, J. Grove, O. A. MacLean, C. Wilkie, G. De Lorenzo, W. Furnon, D. Cantoni, S. Scott, N. Logan, S. Ashraf, M. Manali, A. Szemiel, V. Cowton, E. Vink, W. T. Harvey, C. Davis, P. Asamaphan, K. Smollett, L. Tong, R. Orton, J. Hughes, P. Holland, V. Silva, D. J. Pascall, K. Puxty, A. da Silva Filipe, G. Yebra, S. Shaaban, M. T. G. Holden, R. M. Pinto, R. Gunson, K. Templeton, P. R. Murcia, A. H. Patel, P.

- Klenerman, S. Dunachie, J. Haughney, D. L. Robertson, M. Palmarini, S. Ray, E. C. Thomson, SARS-CoV-2 Omicron is an immune escape variant with an altered cell entry pathway. *Nat. Microbiol.* **7**, 1161–1179 (2022).
7. D. Planas, N. Saunders, P. Maes, F. Guivel-Benhassine, C. Planchais, J. Buchrieser, W.-H. Bolland, F. Porrot, I. Staropoli, F. Lemoine, H. Péré, D. Veyer, J. Puech, J. Rodary, G. Baele, S. Dellicour, J. Raymenants, S. Gorissen, C. Geenen, B. Vanmechelen, T. Wawina-Bokalanga, J. Martí-Carreras, L. Cuypers, A. Sève, L. Hocqueloux, T. Prazuck, F. A. Rey, E. Simon-Loriere, T. Bruel, H. Mouquet, E. André, O. Schwartz, Considerable escape of SARS-CoV-2 Omicron to antibody neutralization. *Nature* **602**, 671–675 (2022).
 8. N. Lasrado, A. Y. Collier, N. P. Hachmann, J. Miller, M. Rowe, E. D. Schonberg, S. L. Rodrigues, A. LaPiana, R. C. Patio, T. Anand, J. Fisher, C. R. Mazurek, R. Guan, K. Wagh, J. Theiler, B. T. Korber, D. H. Barouch, Neutralization escape by SARS-CoV-2 Omicron subvariant BA.2.86. *Vaccine* **41**, 6904–6909 (2023).
 9. A. Wilhelm, M. Widera, K. Grikscheit, T. Toptan, B. Schenk, C. Pallas, M. Metzler, N. Kohmer, S. Hoehl, R. Marschalek, E. Herrmann, F. A. Helfritz, T. Wolf, U. Goetsch, S. Ciesek, Limited neutralisation of the SARS-CoV-2 Omicron subvariants BA.1 and BA.2 by convalescent and vaccine serum and monoclonal antibodies. *eBioMedicine* **82** (2022).
 10. N. Lasrado, A. Y. Collier, J. Miller, N. P. Hachmann, J. Liu, M. Sciacca, C. Wu, T. Anand, E. A. Bondzie, J. L. Fisher, C. R. Mazurek, R. C. Patio, O. Powers, S. L. Rodrigues, M. Rowe, N. Surve, D. M. Ty, B. Korber, D. H. Barouch, Waning Immunity Against XBB.1.5 Following Bivalent mRNA Boosters. *bioRxiv* [Preprint] (2023). <https://doi.org/10.1101/2023.01.22.525079>.
 11. D. Delaune, V. Hul, E. A. Karlsson, A. Hassanin, T. P. Ou, A. Baidaliuk, F. Gámbaro, M. Prot, V. T. Tu, S. Chea, L. Keatts, J. Mazet, C. K. Johnson, P. Buchy, P. Dussart, T. Goldstein, E. Simon-Lorière, V. Duong, A novel SARS-CoV-2 related coronavirus in bats from Cambodia. *Nat. Commun.* **12**, 6563 (2021).
 12. S. Wacharapluesadee, C. W. Tan, P. Maneeorn, P. Duengkae, F. Zhu, Y. Joyjinda, T. Kaewpom, W. N. Chia, W. Ampoot, B. L. Lim, K. Worachotsueptrakun, V. C.-W. Chen, N. Sirichan, C. Ruchisrisarod, A. Rodpan, K. Noradechanon, T. Phaichana, N. Jantararat, B. Thongnumchaima, C. Tu, G. Cramer, M. M. Stokes, T. Hemachudha, L.-F. Wang, Evidence for SARS-CoV-2 related coronaviruses circulating in bats and pangolins in Southeast Asia. *Nat. Commun.* **12**, 972 (2021).

13. N. Zamorano Cuervo, N. Grandvaux, ACE2: Evidence of role as entry receptor for SARS-CoV-2 and implications in comorbidities. *eLife* **9**, e61390 (2020).
14. M. Hoffmann, H. Kleine-Weber, S. Schroeder, N. Krüger, T. Herrler, S. Erichsen, T. S. Schiergens, G. Herrler, N.-H. Wu, A. Nitsche, M. A. Müller, C. Drosten, S. Pöhlmann, SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **181**, 271-280.e8 (2020).
15. C. O. Barnes, C. A. Jette, M. E. Abernathy, K.-M. A. Dam, S. R. Esswein, H. B. Gristick, A. G. Malyutin, N. G. Sharaf, K. E. Huey-Tubman, Y. E. Lee, D. F. Robbiani, M. C. Nussenzweig, A. P. West, P. J. Bjorkman, SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature* **588**, 682–687 (2020).
16. A. J. Greaney, A. N. Loes, L. E. Gentles, K. H. D. Crawford, T. N. Starr, K. D. Malone, H. Y. Chu, J. D. Bloom, Antibodies elicited by mRNA-1273 vaccination bind more broadly to the receptor binding domain than do those from SARS-CoV-2 infection. *Sci. Transl. Med.* **13**, eabi9915 (2021).
17. A. J. Greaney, T. N. Starr, P. Gilchuk, S. J. Zost, E. Binshtein, A. N. Loes, S. K. Hilton, J. Huddleston, R. Eguia, K. H. D. Crawford, A. S. Dingens, R. S. Nargi, R. E. Sutton, N. Suryadevara, P. W. Rothlauf, Z. Liu, S. P. J. Whelan, R. H. Carnahan, J. E. Crowe, J. D. Bloom, Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition. *Cell Host Microbe* **29**, 44-57.e9 (2021).
18. A. J. Greaney, A. N. Loes, K. H. D. Crawford, T. N. Starr, K. D. Malone, H. Y. Chu, J. D. Bloom, Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. *Cell Host Microbe* **29**, 463-476.e6 (2021).
19. A. J. Greaney, T. N. Starr, R. T. Eguia, A. N. Loes, K. Khan, F. Karim, S. Cele, J. E. Bowen, J. K. Logue, D. Corti, D. Veisler, H. Y. Chu, A. Sigal, J. D. Bloom, A SARS-CoV-2 variant elicits an antibody response with a shifted immunodominance hierarchy. *PLOS Pathog.* **18**, e1010248 (2022).
20. A. J. Greaney, R. T. Eguia, T. N. Starr, K. Khan, N. Franko, J. K. Logue, S. M. Lord, C. Speake, H. Y. Chu, A. Sigal, J. D. Bloom, The SARS-CoV-2 Delta variant induces an antibody response largely focused on class 1 and 2 antibody epitopes. *PLOS Pathog.* **18**, e1010592 (2022).

21. C. A. Jette, A. A. Cohen, P. N. P. Gnanapragasam, F. Muecksch, Y. E. Lee, K. E. Huey-Tubman, F. Schmidt, T. Hatzioannou, P. D. Bieniasz, M. C. Nussenzweig, A. P. West, J. R. Keeffe, P. J. Bjorkman, C. O. Barnes, Broad cross-reactivity across sarbecoviruses exhibited by a subset of COVID-19 donor-derived neutralizing antibodies. *Cell Rep.* **36** (2021).
22. A. Ben Chorin, G. Masrati, A. Kessel, A. Narunsky, J. Sprinzak, S. Lahav, H. Ashkenazy, N. Ben-Tal, ConSurf-DB: An accessible repository for the evolutionary conservation patterns of the majority of PDB proteins. *Protein Sci.* **29**, 258–267 (2020).
23. E. Wang, A. K. Chakraborty, Design of immunogens for eliciting antibody responses that may protect against SARS-CoV-2 variants. *PLOS Comput. Biol.* **18**, e1010563 (2022).
24. K. D. Brune, D. B. Leneghan, I. J. Brian, A. S. Ishizuka, M. F. Bachmann, S. J. Draper, S. Biswas, M. Howarth, Plug-and-Display: decoration of Virus-Like Particles via isopeptide bonds for modular immunization. *Sci. Rep.* **6**, 19234 (2016).
25. B. Zakeri, J. O. Fierer, E. Celik, E. C. Chittock, U. Schwarz-Linek, V. T. Moy, M. Howarth, Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proc. Natl. Acad. Sci.* **109**, E690–E697 (2012).
26. A. A. Cohen, P. N. P. Gnanapragasam, Y. E. Lee, P. R. Hoffman, S. Ou, L. M. Kakutani, J. R. Keeffe, H.-J. Wu, M. Howarth, A. P. West, C. O. Barnes, M. C. Nussenzweig, P. J. Bjorkman, Mosaic nanoparticles elicit cross-reactive immune responses to zoonotic coronaviruses in mice. *Science* **371**, 735–741 (2021).
27. A. A. Cohen, N. van Doremalen, A. J. Greaney, H. Andersen, A. Sharma, T. N. Starr, J. R. Keeffe, C. Fan, J. E. Schulz, P. N. P. Gnanapragasam, L. M. Kakutani, A. P. West, G. Saturday, Y. E. Lee, H. Gao, C. A. Jette, M. G. Lewis, T. K. Tan, A. R. Townsend, J. D. Bloom, V. J. Munster, P. J. Bjorkman, Mosaic RBD nanoparticles protect against challenge by diverse sarbecoviruses in animal models. *Science* **377**, eabq0839 (2022).
28. W. Dong, H. Mead, L. Tian, J.-G. Park, J. I. Garcia, S. Jaramillo, T. Barr, D. S. Kollath, V. K. Coyne, N. E. Stone, A. Jones, J. Zhang, A. Li, L.-S. Wang, M. Milanes-Yearsley, J. B. Torrelles, L. Martinez-Sobrido, P. S. Keim, B. M. Barker, M. A. Caligiuri, J. Yu, The K18-Human ACE2 Transgenic Mouse Model Recapitulates Non-severe and Severe COVID-19 in Response to an Infectious Dose of the SARS-CoV-2 Virus. *J. Virol.* **96**, e00964-21 (2022).

29. T. N. Starr, A. J. Greaney, S. K. Hilton, D. Ellis, K. H. D. Crawford, A. S. Dingens, M. J. Navarro, J. E. Bowen, M. A. Tortorici, A. C. Walls, N. P. King, D. Veelsler, J. D. Bloom, Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell* **182**, 1295-1310.e20 (2020).
30. J. Dong, S. J. Zost, A. J. Greaney, T. N. Starr, A. S. Dingens, E. C. Chen, R. E. Chen, J. B. Case, R. E. Sutton, P. Gilchuk, J. Rodriguez, E. Armstrong, C. Gainza, R. S. Nargi, E. Binshtein, X. Xie, X. Zhang, P.-Y. Shi, J. Logue, S. Weston, M. E. McGrath, M. B. Frieman, T. Brady, K. M. Tuffy, H. Bright, Y.-M. Loo, P. M. McTamney, M. T. Esser, R. H. Carnahan, M. S. Diamond, J. D. Bloom, J. E. Crowe, Genetic and structural basis for SARS-CoV-2 variant neutralization by a two-antibody cocktail. *Nat. Microbiol.* **6**, 1233–1244 (2021).
31. A. J. Greaney, T. N. Starr, C. O. Barnes, Y. Weisblum, F. Schmidt, M. Caskey, C. Gaebler, A. Cho, M. Agudelo, S. Finkin, Z. Wang, D. Poston, F. Muecksch, T. Hatziioannou, P. D. Bieniasz, D. F. Robbiani, M. C. Nussenzweig, P. J. Bjorkman, J. D. Bloom, Mapping mutations to the SARS-CoV-2 RBD that escape binding by different classes of antibodies. *Nat. Commun.* **12**, 4196 (2021).
32. T. N. Starr, A. J. Greaney, A. S. Dingens, J. D. Bloom, Complete map of SARS-CoV-2 RBD mutations that escape the monoclonal antibody LY-CoV555 and its cocktail with LY-CoV016. *Cell Rep. Med.* **2** (2021).
33. T. N. Starr, N. Czudnochowski, Z. Liu, F. Zatta, Y.-J. Park, A. Addetia, D. Pinto, M. Beltramello, P. Hernandez, A. J. Greaney, R. Marzi, W. G. Glass, I. Zhang, A. S. Dingens, J. E. Bowen, M. A. Tortorici, A. C. Walls, J. A. Wojcechowskyj, A. De Marco, L. E. Rosen, J. Zhou, M. Montiel-Ruiz, H. Kaiser, J. R. Dillen, H. Tucker, J. Bassi, C. Silacci-Fregni, M. P. Housley, J. di Iulio, G. Lombardo, M. Agostini, N. Sprugasci, K. Culap, S. Jaconi, M. Meury, E. Dellota Jr, R. Abdelnabi, S.-Y. C. Foo, E. Cameroni, S. Stumpf, T. I. Croll, J. C. Nix, C. Havenar-Daughton, L. Piccoli, F. Benigni, J. Neyts, A. Telenti, F. A. Lempp, M. S. Pizzuto, J. D. Chodera, C. M. Hebner, H. W. Virgin, S. P. J. Whelan, D. Veelsler, D. Corti, J. D. Bloom, G. Snell, SARS-CoV-2 RBD antibodies that maximize breadth and resistance to escape. *Nature* **597**, 97–102 (2021).
34. L. Yang, M. V. Beek, Z. Wang, F. Muecksch, M. Canis, T. Hatziioannou, P. D. Bieniasz, M. C. Nussenzweig, A. K. Chakraborty, Antigen presentation dynamics shape the antibody response to variants like SARS-CoV-2 Omicron after multiple vaccinations with the original strain. *Cell Rep.* **42** (2023).

35. E. Andreano, I. Paciello, G. Pierleoni, G. Piccini, V. Abbiento, G. Antonelli, P. Pileri, N. Manganaro, E. Pantano, G. Maccari, S. Marchese, L. Donnici, L. Benincasa, G. Giglioli, M. Leonardi, C. De Santi, M. Fabbiani, I. Rancan, M. Tumbarello, F. Montagnani, C. Sala, D. Medini, R. De Francesco, E. Montomoli, R. Rappuoli, B cell analyses after SARS-CoV-2 mRNA third vaccination reveals a hybrid immunity like antibody response. *Nat. Commun.* **14**, 53 (2023).
36. A. Kuriata, V. Iglesias, M. Kurcinski, S. Ventura, S. Kmiecik, Aggrescan3D standalone package for structure-based prediction of protein aggregation properties. *Bioinformatics* **35**, 3834–3835 (2019).
37. C. E. Shannon, A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423 (1948).
38. M. Letko, A. Marzi, V. Munster, Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nat. Microbiol.* **5**, 562–569 (2020).
39. E. W. Sayers, E. E. Bolton, J. R. Brister, K. Canese, J. Chan, D. C. Comeau, R. Connor, K. Funk, C. Kelly, S. Kim, T. Madej, A. Marchler-Bauer, C. Lanczycki, S. Lathrop, Z. Lu, F. Thibaud-Nissen, T. Murphy, L. Phan, Y. Skripchenko, T. Tse, J. Wang, R. Williams, B. W. Trawick, K. D. Pruitt, S. T. Sherry, Database resources of the national center for biotechnology information. *Nucleic Acids Res.* **50**, D20–D26 (2022).
40. S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).
41. J. D. Thompson, D. G. Higgins, T. J. Gibson, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680 (1994).
42. J. F. Scheid, C. O. Barnes, B. Eraslan, A. Hudak, J. R. Keeffe, L. A. Cosimi, E. M. Brown, F. Muecksch, Y. Weisblum, S. Zhang, T. Delorey, A. E. Woolley, F. Ghantous, S.-M. Park, D. Phillips, B. Tusi, K. E. Huey-Tubman, A. A. Cohen, P. N. P. Gnanapragasam, K. Rzasa, T. Hatziioanno, M. A. Durney, X. Gu, T. Tada, N. R. Landau, A. P. West, O. Rozenblatt-Rosen, M. S. Seaman, L. R. Baden, D. B. Graham, J. Deguine, P. D. Bieniasz, A. Regev, D. Hung, P. J. Bjorkman, R. J. Xavier, B cell genomics behind cross-neutralization of SARS-CoV-2 variants and SARS-CoV. *Cell* **184**, 3205–3221.e24 (2021).

43. A. A. Cohen, J. R. Keeffe, A. Schiepers, S. E. Dross, A. J. Greaney, A. V. Rorick, H. Gao, P. N. P. Gnanapragasam, C. Fan, A. P. West, A. I. Ramsingh, J. H. Erasmus, J. D. Pata, H. Muramatsu, N. Pardi, P. J. C. Lin, S. Baxter, R. Cruz, M. Quintanar-Audelo, E. Robb, C. Serrano-Amatriain, L. Magneschi, I. G. Fotheringham, D. H. Fuller, G. D. Victora, P. J. Bjorkman, Mosaic sarbecovirus nanoparticles elicit cross-reactive responses in pre-vaccinated animals. *Cell* **187**, 5554-5571.e19 (2024).
44. A. A. Cohen, Z. Yang, P. N. P. Gnanapragasam, S. Ou, K.-M. A. Dam, H. Wang, P. J. Bjorkman, Construction, characterization, and immunization of nanoparticles that display a diverse array of influenza HA trimers. *PLOS ONE* **16**, e0247963 (2021).
45. D. Goldblatt, G. Alter, S. Crotty, S. A. Plotkin, Correlates of protection against SARS-CoV-2 infection and COVID-19 disease. *Immunol. Rev.* **310**, 6–26 (2022).
46. K. S. Corbett, M. C. Nason, B. Flach, M. Gagne, S. O’Connell, T. S. Johnston, S. N. Shah, V. V. Edara, K. Floyd, L. Lai, C. McDanal, J. R. Francica, B. Flynn, K. Wu, A. Choi, M. Koch, O. M. Abiona, A. P. Werner, J. I. Moliva, S. F. Andrew, M. M. Donaldson, J. Fintzi, D. R. Flebbe, E. Lamb, A. T. Noe, S. T. Nurmukhambetova, S. J. Provost, A. Cook, A. Dodson, A. Faudree, J. Greenhouse, S. Kar, L. Pessaint, M. Porto, K. Steingrebe, D. Valentin, S. Zouantcha, K. W. Bock, M. Minai, B. M. Nagata, R. van de Wetering, S. Boyoglu-Barnum, K. Leung, W. Shi, E. S. Yang, Y. Zhang, J.-P. M. Todd, L. Wang, G. S. Alvarado, H. Andersen, K. E. Foulds, D. K. Edwards, J. R. Mascola, I. N. Moore, M. G. Lewis, A. Carfi, D. Montefiori, M. S. Suthar, A. McDermott, M. Roederer, N. J. Sullivan, D. C. Douek, B. S. Graham, R. A. Seder, Immune correlates of protection by mRNA-1273 vaccine against SARS-CoV-2 in nonhuman primates. *Science* **373**, eabj0299 (2021).
47. P. B. Gilbert, D. C. Montefiori, A. B. McDermott, Y. Fong, D. Benkeser, W. Deng, H. Zhou, C. R. Houchens, K. Martins, L. Jayashankar, F. Castellino, B. Flach, B. C. Lin, S. O’Connell, C. McDanal, A. Eaton, M. Sarzotti-Kelsoe, Y. Lu, C. Yu, B. Borate, L. W. P. van der Laan, N. S. Hejazi, C. Huynh, J. Miller, H. M. El Sahly, L. R. Baden, M. Baron, L. De La Cruz, C. Gay, S. Kalams, C. F. Kelley, M. P. Andrasik, J. G. Kublin, L. Corey, K. M. Neuzil, L. N. Carpp, R. Pajon, D. Follmann, R. O. Donis, R. A. Koup, Immune Assays Team, Moderna, Inc. Team, Coronavirus Vaccine Prevention Network (CoVPN)/Coronavirus Efficacy (COVE) Team, UNITED STATES GOVERNMENT (USG)/COVPN BIOSTATISTICS TEAM, Immune correlates analysis of the mRNA-1273 COVID-19 vaccine efficacy clinical trial. *Science* **375**, 43–50 (2022).
48. A. Amitai, M. Sangesland, R. M. Barnes, D. Rohrer, N. Lonberg, D. Lingwood, A. K. Chakraborty, Defining and Manipulating B Cell Immunodominance Hierarchies to

- Elicit Broadly Neutralizing Antibody Responses against Influenza Virus. *Cell Syst.* **11**, 573-588.e9 (2020).
49. Z. Mu, K. Wiehe, K. O. Saunders, R. Henderson, D. W. Cain, R. Parks, D. Martik, K. Mansouri, R. J. Edwards, A. Newman, X. Lu, S.-M. Xia, A. Eaton, M. Bonsignori, D. Montefiori, Q. Han, S. Venkatayogi, T. Evangelous, Y. Wang, W. Rountree, B. Korber, K. Wagh, Y. Tam, C. Barbosa, S. M. Alam, W. B. Williams, M. Tian, F. W. Alt, N. Pardi, D. Weissman, B. F. Haynes, mRNA-encoded HIV-1 Env trimer ferritin nanoparticles induce monoclonal antibodies that neutralize heterologous HIV-1 isolates in mice. *Cell Rep.* **38** (2022).
 50. G. Ueda, A. Antanasijevic, J. A. Fallas, W. Sheffler, J. Copps, D. Ellis, G. B. Hutchinson, A. Moyer, A. Yasmeen, Y. Tsybovsky, Y.-J. Park, M. J. Bick, B. Sankaran, R. A. Gillespie, P. J. Brouwer, P. H. Zwart, D. Veessler, M. Kanekiyo, B. S. Graham, R. W. Sanders, J. P. Moore, P. J. Klasse, A. B. Ward, N. P. King, D. Baker, Tailored design of protein nanoparticle scaffolds for multivalent presentation of viral glycoprotein antigens. *eLife* **9**, e57659 (2020).
 51. H. B. Gristick, H. Hartweger, M. Loewe, J. van Schooten, V. Ramos, T. Y. Oliveira, Y. Nishimura, N. S. Koranda, A. Wall, K.-H. Yao, D. Poston, A. Gazumyan, M. Wiatr, M. Horning, J. R. Keefe, M. A. G. Hoffmann, Z. Yang, M. E. Abernathy, K.-M. A. Dam, H. Gao, P. N. P. Gnanapragasam, L. M. Kakutani, A. J. Pavlovitch-Bedzyk, M. S. Seaman, M. Howarth, A. T. McGuire, L. Stamatatos, M. A. Martin, A. P. West, M. C. Nussenzweig, P. J. Bjorkman, CD4 binding site immunogens elicit heterologous anti-HIV-1 neutralizing antibodies in transgenic and wild-type animals. *Sci. Immunol.* **8**, eade6364 (2023).
 52. W. Wang, B. Huang, Y. Zhu, W. Tan, M. Zhu, Ferritin nanoparticle-based SARS-CoV-2 RBD vaccine induces a persistent antibody response and long-term memory in mice. *Cell. Mol. Immunol.* **18**, 749–751 (2021).
 53. M. G. Joyce, W.-H. Chen, R. S. Sankhala, A. Hajduczki, P. V. Thomas, M. Choe, E. J. Martinez, W. C. Chang, C. E. Peterson, E. B. Morrison, C. Smith, R. E. Chen, A. Ahmed, L. Wiczorek, A. Anderson, J. B. Case, Y. Li, T. Oertel, L. Rosado, A. Ganesh, C. Whalen, J. M. Carmen, L. Mendez-Rivera, C. P. Karch, N. Gohain, Z. Villar, D. McCurdy, Z. Beck, J. Kim, S. Shrivastava, O. Jobe, V. Dussupt, S. Molnar, U. Tran, C. B. Kannadka, S. Soman, C. Kuklis, M. Zemil, H. Khanh, W. Wu, M. A. Cole, D. K. Duso, L. W. Kummer, T. J. Lang, S. E. Muncil, J. R. Currier, S. J. Krebs, V. R. Polonis, S. Rajan, P. M. McTamney, M. T. Esser, W. W. Reiley, M. Rolland, N. de Val, M. S. Diamond, G. D. Gromowski, G. R. Matyas, M. Rao, N. L. Michael, K. Modjarrad,

- SARS-CoV-2 ferritin nanoparticle vaccines elicit broad SARS coronavirus immunogenicity. *Cell Rep.* **37** (2021).
54. Y. Zhang, J. Sun, J. Zheng, S. Li, H. Rao, J. Dai, Z. Zhang, Y. Wang, D. Liu, Z. Chen, W. Ran, A. Zhu, F. Li, Q. Yan, Y. Wang, K. Yu, S. Zhang, D. Wang, Y. Tang, B. Liu, L. Cheng, J. Huo, S. Perlman, J. Zhao, J. Zhao, Mosaic RBD Nanoparticles Elicit Protective Immunity Against Multiple Human Coronaviruses in Animal Models. *Adv. Sci.* **11**, 2303366 (2024).
 55. M. G. Joyce, H. A. D. King, I. Elakhal-Naouar, A. Ahmed, K. K. Peachman, C. Macedo Cincotta, C. Subra, R. E. Chen, P. V. Thomas, W.-H. Chen, R. S. Sankhala, A. Hajduczki, E. J. Martinez, C. E. Peterson, W. C. Chang, M. Choe, C. Smith, P. J. Lee, J. A. Headley, M. G. Taddese, H. A. Elyard, A. Cook, A. Anderson, K. McGuckin Wuertz, M. Dong, I. Swafford, J. B. Case, J. R. Currier, K. G. Lal, S. Molnar, M. S. Nair, V. Dussupt, S. P. Daye, X. Zeng, E. K. Barkei, H. M. Staples, K. Alfson, R. Carrion, S. J. Krebs, D. Paquin-Proulx, N. Karasavva, V. R. Polonis, L. L. Jagodzinski, M. F. Amare, S. Vasan, P. T. Scott, Y. Huang, D. D. Ho, N. de Val, M. S. Diamond, M. G. Lewis, M. Rao, G. R. Matyas, G. D. Gromowski, S. A. Peel, N. L. Michael, D. L. Bolton, K. Modjarrad, A SARS-CoV-2 ferritin nanoparticle vaccine elicits protective immune responses in nonhuman primates. *Sci. Transl. Med.* **14**, eabi5735 (2021).
 56. P. A.-B. Weidenbacher, M. Sanyal, N. Friedland, S. Tang, P. S. Arunachalam, M. Hu, O. S. Kumru, M. K. Morris, J. Fontenot, L. Shirreff, J. Do, Y.-C. Cheng, G. Vasudevan, M. B. Feinberg, F. J. Villinger, C. Hanson, S. B. Joshi, D. B. Volkin, B. Pulendran, P. S. Kim, A ferritin-based COVID-19 nanoparticle vaccine that elicits robust, durable, broad-spectrum neutralizing antisera in non-human primates. *Nat. Commun.* **14**, 2149 (2023).
 57. S. C. Johnston, K. M. Ricks, I. Lakhal-Naouar, A. Jay, C. Subra, J. L. Raymond, H. A. D. King, F. Rossi, T. L. Clements, D. Fetterer, S. Tostenson, C. M. Cincotta, H. R. Hack, C. Kuklis, S. Soman, J. King, K. K. Peachman, D. Kim, W.-H. Chen, R. S. Sankhala, E. J. Martinez, A. Hajduczki, W. C. Chang, M. Choe, P. V. Thomas, C. E. Peterson, A. Anderson, I. Swafford, J. R. Currier, D. Paquin-Proulx, L. L. Jagodzinski, G. R. Matyas, M. Rao, G. D. Gromowski, S. A. Peel, L. White, J. M. Smith, J. W. Hooper, N. L. Michael, K. Modjarrad, M. G. Joyce, A. Nalca, D. L. Bolton, M. L. M. Pitt, A SARS-CoV-2 Spike Ferritin Nanoparticle Vaccine Is Protective and Promotes a Strong Immunological Response in the Cynomolgus Macaque Coronavirus Disease 2019 (COVID-19) Model. *Vaccines* **10**, 717 (2022).

58. R. S. Sankhala, K. G. Lal, J. L. Jensen, V. Dussupt, L. Mendez-Rivera, H. Bai, L. Wiczorek, S. V. Mayer, M. Zemil, D. A. Wagner, S. M. Townsley, A. Hajducski, W. C. Chang, W.-H. Chen, G. C. Donofrio, N. Jian, H. A. D. King, C. G. Lorang, E. J. Martinez, P. A. Rees, C. E. Peterson, F. Schmidt, T. J. Hart, D. K. Duso, L. W. Kummer, S. P. Casey, J. K. Williams, S. Kannan, B. M. Slike, L. Smith, I. Swafford, P. V. Thomas, U. Tran, J. R. Currier, D. L. Bolton, E. Davidson, B. J. Doranz, T. Hatzioannou, P. D. Bieniasz, D. Paquin-Proulx, W. W. Reiley, M. Rolland, N. J. Sullivan, S. Vasan, N. D. Collins, K. Modjarrad, G. D. Gromowski, V. R. Polonis, N. L. Michael, S. J. Krebs, M. G. Joyce, Diverse array of neutralizing antibodies elicited upon Spike Ferritin Nanoparticle vaccination in rhesus macaques. *Nat. Commun.* **15**, 200 (2024).
59. J. Yu, P. V. Thomas, M. Sciacca, C. Wu, J. Liu, X. He, J. Miller, N. P. Hachmann, N. Surve, K. McMahan, C. Jacob-Dolan, O. Powers, K. Hall, J. Barrett, D. Hope, C. R. Mazurek, T. Murdza, W. C. Chang, E. Golub, P. A. Rees, C. E. Peterson, A. Hajducski, W.-H. Chen, E. J. Martinez, E. Hussin, C. Lange, H. Gong, G. R. Matyas, M. Rao, M. Suthar, M. Boursiquot, A. Cook, L. Pessaint, M. G. Lewis, H. Andersen, D. L. Bolton, N. L. Michael, M. G. Joyce, K. Modjarrad, D. H. Barouch, Ad26.COVS.2 and SARS-CoV-2 spike protein ferritin nanoparticle vaccine protect against SARS-CoV-2 Omicron BA.5 challenge in macaques. *Cell Rep. Med.* **4** (2023).
60. A. E. Powell, K. Zhang, M. Sanyal, S. Tang, P. A. Weidenbacher, S. Li, T. D. Pham, J. E. Pak, W. Chiu, P. S. Kim, A Single Immunization with Spike-Functionalized Ferritin Vaccines Elicits Neutralizing Antibody Responses against SARS-CoV-2 in Mice. *ACS Cent. Sci.* **7**, 183–199 (2021).
61. Z.-H. Liu, Z.-F. Deng, Y. Lu, W.-H. Fang, F. He, A modular and self-adjuvanted multivalent vaccine platform based on porcine circovirus virus-like nanoparticles. *J. Nanobiotechnology* **20**, 493 (2022).
62. V. Lampinen, S. Gröhn, S. Soppela, V. Blazevic, V. P. Hytönen, M. M. Hankaniemi, SpyTag/SpyCatcher display of influenza M2e peptide on norovirus-like particle provides stronger immunization than direct genetic fusion. *Front. Cell. Infect. Microbiol.* **13** (2023).
63. P. Ximba, R. Chapman, A. Meyers, E. Margolin, M. T. van Diepen, A. F. Sander, J. Woodward, P. L. Moore, A.-L. Williamson, E. P. Rybicki, Development of a synthetic nanoparticle vaccine presenting the HIV-1 envelope glycoprotein. *Nanotechnology* **33**, 485102 (2022).

64. J. M. J. Tas, J.-H. Koo, Y.-C. Lin, Z. Xie, J. M. Steichen, A. M. Jackson, B. M. Hauser, X. Wang, C. A. Cottrell, J. L. Torres, J. E. Warner, K. H. Kirsch, S. R. Weldon, B. Groschel, B. Nogal, G. Ozorowski, S. Bangaru, N. Phelps, Y. Adachi, S. Eskandarzadeh, M. Kubitz, D. R. Burton, D. Lingwood, A. G. Schmidt, U. Nair, A. B. Ward, W. R. Schief, F. D. Batista, Antibodies from primary humoral responses modulate the recruitment of naive B cells during secondary responses. *Immunity* **55**, 1856-1871.e6 (2022).
65. D. Schaefer-Babajew, Z. Wang, F. Muecksch, A. Cho, M. Loewe, M. Cipolla, R. Raspe, B. Johnson, M. Canis, J. DaSilva, V. Ramos, M. Turroja, K. G. Millard, F. Schmidt, L. Witte, J. Dizon, I. Shimeliovich, K.-H. Yao, T. Y. Oliveira, A. Gazumyan, C. Gaebler, P. D. Bieniasz, T. Hatzioannou, M. Caskey, M. C. Nussenzweig, Antibody feedback regulates immune memory after SARS-CoV-2 mRNA vaccination. *Nature* **613**, 735–742 (2023).
66. A. Schiepers, M. F. L. van 't Wout, A. J. Greaney, T. Zang, H. Muramatsu, P. J. C. Lin, Y. K. Tam, L. Mesin, T. N. Starr, P. D. Bieniasz, N. Pardi, J. D. Bloom, G. D. Vitoria, Molecular fate-mapping of serum antibody responses to repeat immunization. *Nature* **615**, 482–489 (2023).
67. H. A. McNamara, A. H. Idris, H. J. Sutton, R. Vistein, B. J. Flynn, Y. Cai, K. Wiehe, K. E. Lyke, D. Chatterjee, N. Kc, S. Chakravarty, B. K. L. Sim, S. L. Hoffman, M. Bonsignori, R. A. Seder, I. A. Cockburn, Antibody Feedback Limits the Expansion of B Cell Responses to Malaria Vaccination but Drives Diversification of the Humoral Response. *Cell Host Microbe* **28**, 572-585.e7 (2020).
68. B. Heyman, Regulation of Antibody Responses via Antibodies, Complement, and Fc Receptors. *Annu. Rev. Immunol.* **18**, 709–737 (2000).
69. M. Kanekiyo, M. G. Joyce, R. A. Gillespie, J. R. Gallagher, S. F. Andrews, H. M. Yassine, A. K. Wheatley, B. E. Fisher, D. R. Ambrozak, A. Creanga, K. Leung, E. S. Yang, S. Boyoglu-Barnum, I. S. Georgiev, Y. Tsybovsky, M. S. Prabhakaran, H. Andersen, W.-P. Kong, U. Baxa, K. L. Zephir, J. E. Ledgerwood, R. A. Koup, P. D. Kwong, A. K. Harris, A. B. McDermott, J. R. Mascola, B. S. Graham, Mosaic nanoparticle display of diverse influenza virus hemagglutinins elicits broad B cell responses. *Nat. Immunol.* **20**, 362–372 (2019).
70. C. Fan, A. A. Cohen, M. Park, A. F.-H. Hung, J. R. Keeffe, P. N. P. Gnanapragasam, Y. E. Lee, H. Gao, L. M. Kakutani, Z. Wu, H. Kleanthous, K. E. Malecek, J. C. Williams, P. J. Bjorkman, Neutralizing monoclonal antibodies elicited by mosaic RBD

- nanoparticles bind conserved sarbecovirus epitopes. *Immunity* **55**, 2419–2435.e10 (2022).
71. F. Schmidt, Y. Weisblum, F. Muecksch, H.-H. Hoffmann, E. Michailidis, J. C. C. Lorenzi, P. Mendoza, M. Rutkowska, E. Bednarski, C. Gaebler, M. Agudelo, A. Cho, Z. Wang, A. Gazumyan, M. Cipolla, M. Caskey, D. F. Robbani, M. C. Nussenzweig, C. M. Rice, T. Hatzioannou, P. D. Bieniasz, Measuring SARS-CoV-2 neutralizing antibody activity using pseudotyped and chimeric viruses. *J. Exp. Med.* **217**, e20201181 (2020).
 72. C.-L. Hsieh, J. A. Goldsmith, J. M. Schaub, A. M. DiVenere, H.-C. Kuo, K. Javanmardi, K. C. Le, D. Wrapp, A. G. Lee, Y. Liu, C.-W. Chou, P. O. Byrne, C. K. Hjorth, N. V. Johnson, J. Ludes-Meyers, A. W. Nguyen, J. Park, N. Wang, D. Amengor, J. J. Lavinder, G. C. Ippolito, J. A. Maynard, I. J. Finkelstein, J. S. McLellan, Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science* **369**, 1501–1505 (2020).
 73. Y. Cao, F. Jian, J. Wang, Y. Yu, W. Song, A. Yisimayi, J. Wang, R. An, X. Chen, N. Zhang, Y. Wang, P. Wang, L. Zhao, H. Sun, L. Yu, S. Yang, X. Niu, T. Xiao, Q. Gu, F. Shao, X. Hao, Y. Xu, R. Jin, Z. Shen, Y. Wang, X. S. Xie, Imprinted SARS-CoV-2 humoral immunity induces convergent Omicron RBD evolution. *Nature* **614**, 521–529 (2023).
 74. A. H. Keeble, P. Turkki, S. Stokes, I. N. A. Khairil Anuar, R. Rahikainen, V. P. Hytönen, M. Howarth, Approaching infinite affinity through engineering of peptide–protein interaction. *Proc. Natl. Acad. Sci.* **116**, 26523–26533 (2019).
 75. A. P. West, L. Scharf, J. Horwitz, F. Klein, M. C. Nussenzweig, P. J. Bjorkman, Computational analysis of anti-HIV-1 antibody neutralization panel data to identify potential functional epitope residues. *Proc. Natl. Acad. Sci.* **110**, 10598–10603 (2013).
 76. R. A. Hills, T. K. Tan, A. A. Cohen, J. R. Keeffe, A. H. Keeble, P. N. P. Gnanaprasam, K. N. Storm, A. V. Rorick, A. P. West, M. L. Hill, S. Liu, J. Gilbert-Jaramillo, M. Afzal, A. Napier, G. Admans, W. S. James, P. J. Bjorkman, A. R. Townsend, M. R. Howarth, Proactive vaccination using multiviral Quartet Nanocages to elicit broad anti-coronavirus responses. *Nat. Nanotechnol.* **19**, 1216–1223 (2024).

STABILIZED GP120-SPECIFIC CD4 FOR NEXT-GENERATION HIV-1 INHIBITORS

A. J. Bahn-Suh, **L. F. Caldera**, P. N. P. Gnanapragasam, J. R. Keeffe, M. S. Seaman, P. J. Bjorkman, S. L. Mayo, Stabilized gp120-specific CD4 for next-generation HIV-1 inhibitors. *bioRxiv* [Preprint] (2026). <https://doi.org/10.64898/2026.03.24.713825>.

3.1 Summary

HIV-1 Env's gp120 subunit uses the T-cell coreceptor CD4 to enter host cells in a manner that prevents the evolution of host resistance by sharing the binding epitope with the footprint of CD4's natural ligands, class II MHC proteins (1, 2). Consequently, CD4-containing biologics, such as CD4-Ig (3, 4) and derivatives (5–9), benefit from this conserved relationship and are promising broad-acting anti-HIV-1 agents that are resistant to viral mutational escape (10). However, these biologics suffer from short serum half-lives in humans (11, 12) and animals (3, 13), likely due to CD4's poor thermostability (14) and/or off-target class II MHC binding (15). This latter property also warrants caution for CD4-containing biologics that could indiscriminately recruit Fc-dependent effector functions against uninfected cells and/or compete with host CD4 for class II MHC during T cell interactions with antigen presenting cells. Here, we describe gp120-specific CD4 (gCD4), which exhibits enhanced thermostability and retains Env, but not class II MHC, binding. CD4-Ig variants incorporating gCD4 did not bind class II MHC on human B cells, displayed greater longevity in human tonsil organoid cultures, showed half-lives equivalent to therapeutic IgG antibodies in mice, and neutralized HIV-1 more broadly and potently compared to the original CD4-Ig molecules. Encouragingly, one variant neutralized 100% of a panel of clinically-relevant HIV-1 strains at titers correlating to infection prevention in humans, outperforming known broadly neutralizing antibodies (16, 17). Thus, gCD4 holds promise for the development of new CD4-containing biologics with best-in-class specificity, pharmacokinetic properties, and neutralization breadth and potency.

3.2 Introduction

Broadly neutralizing antibodies (bNAbs) are clinically effective against HIV-1 but are susceptible to existing viral resistance or the development of viral mutational escape (16, 18, 19). Addressing this shortfall with a pan-neutralizing and potent antibody that is resilient to HIV-1 escape has long been sought for therapeutic use. The pursuit of this goal led to the development of CD4-containing biologics such as CD4-Ig (3, 4), eCD4-Ig (7), PRO-542 (6), and others (5, 8, 9, 20). These biologics fuse the two N-terminal domains of human CD4 (D1D2) to IgG Fc in various permutations and are attractive for therapeutic development because HIV-1 uses the gp120 subunit of its trimeric spike protein, Env, to bind to host cell CD4 and chemokine coreceptor to accomplish viral entry (21–25). HIV-1 does this by first engaging CD4 through Env's gp120 subunits, which then initiate conformational changes in Env to facilitate subsequent coreceptor binding and ultimately membrane fusion (26, 27). Therefore, a molecule such as eCD4-Ig, which contains two copies of CD4 and chemokine coreceptor mimetic peptides, is a broad neutralizer that is resilient to and constrains mutational escape—if HIV-1 were to evade neutralization by eCD4-Ig through mutations that reduce binding to CD4 or chemokine coreceptor, these mutations would also reduce HIV-1's fitness. Evidence for this restrictive mechanism was demonstrated when viral swarms of HIV-1 under selective pressure by eCD4-Ig were unable to fully escape neutralization, and any viruses with partial resistance to eCD4-Ig exhibited markedly reduced infectivity (10).

Various studies have shown CD4-containing biologics to prophylactically protect from HIV-1 infection and reduce viral loads during active infection *in vivo*, but these studies also showed evidence of poor pharmacokinetics (11, 12, 15). For example, CD4-Ig and PRO-542 display half-lives of 2-4 days in mice (14) and humans (11, 12, 15), while typical IgGs have half-lives of 2-3 weeks in humans (28). Many factors can affect serum half-life, but one property of CD4 that may be relevant to its pharmacokinetics is its low thermostability. Therapeutic monoclonal antibodies exhibit apparent melting temperatures (T_{ms}) ranging from 60-90°C (29), but the CD4 domains used in CD4-containing biologics have much lower T_{ms} (**Extended Data Table 1**). The poor thermostability of soluble CD4 constructs could

lead to a greater propensity for the denaturation, aggregation, or misfolding of CD4 and contribute to short half-lives *in vivo*.

In addition, target-mediated drug disposition (15, 30) may play a role in the short serum half-lives of CD4-containing biologics. The natural ligands for CD4 are class II MHC proteins (31) (MHC II), which are highly expressed on the surfaces of antigen-presenting cells (APCs) such as dendritic cells, macrophages, and B cells (32). In a recent study relevant to the pharmacokinetics of MHC II-binding proteins, two T cell receptor (TCR) like antibodies that either bound to all or to select peptide-MHC II complexes were investigated for their pharmacokinetic properties (33). The antibody that bound to all peptide-MHC II exhibited rapid clearance, whereas the other exhibited more typical antibody pharmacokinetics. In a similar manner to the TCR-like antibodies, CD4-Ig may experience accelerated clearance through binding to MHC II.

We hypothesized that addressing these issues by stabilizing soluble CD4 and engineering it to preferentially bind HIV-1 gp120 and not MHC II would improve the pharmacokinetics of CD4-containing biologics. However, gp120 and MHC II share nearly identical binding surfaces on CD4—a feature that likely helps HIV-1 avoid the natural selection of protective CD4 polymorphisms (34) as CD4 must maintain the ability to bind MHC II for a functional adaptive immune system. While beneficial for HIV-1, the shared binding interface on CD4 presented a considerable obstacle to our endeavor. Taking inspiration from the “divide and conquer” paradigm in software engineering (exemplified by the merge-sort algorithm (35)), we used computational and rational design methods to engineer CD4 variants that individually possess desired attributes before merging their properties in a final construct, which we refer to as gCD4 (for gp120-specific CD4). Compared to CD4-Ig molecules containing wild-type CD4, gCD4-Ig variants demonstrate better pharmacokinetics in tonsil organoid and humanized mice models and are more potent HIV-1 neutralizers, suggesting that the incorporation of gCD4 in future CD4-containing biologics would enhance their therapeutic properties.

3.3 Results

3.3.1 Electrostatic repulsions abrogated MHC II binding

We analyzed the co-crystal structures of the HIV-1 gp120- and human MHC II-CD4 complexes (1, 2) in PyMol (36) to identify residues that could be substituted to selectively abrogate MHC II binding. The gp120-CD4 interface spans amino acid positions 23-64 in CD4, while the MHC II-CD4 interface spans amino acid positions 35-63 in CD4, a subset of the gp120 binding site (**Figure 1a**). A critical difference between the two complexes is that gp120 is a monomer while MHC II is a heterodimer composed of MHC II α and MHC II β . We aimed to exploit this difference by identifying CD4 residues unessential for binding gp120 and simultaneously crucial for binding MHC II α or MHC II β . This analysis led to four observations. First, there are up to eight MHC II β interactions but only three MHC II α interactions: CD4 residues R59, S60, and D63 interact with the corresponding MHC II α chain residues T90, E88, and K176. Targeting the fewer number of MHC II α interactions would be more feasible than targeting the MHC II β interactions. Second, MHC haplotypes are naturally diverse, but MHC II α is invariant at these positions, circumventing the need to account for polymorphisms in this interface (**Figure 1b**). Third, CD4 S60 and D63 are located at the periphery of the CD4-gp120 interface and have been shown in previous alanine-scanning experiments (37) to be dispensable for gp120 binding. Therefore, mutating these residues should have minimal effects on CD4's affinity for gp120 and should preserve the ability of CD4-containing biologics to restrict HIV-1 mutational escape. Finally, these positions are near the N- and C-termini of a short α -helix in CD4. An α -helix possesses a macrodipole that is more positively charged near its N-terminus and more negatively charged near its C-terminus. Thus, offsetting this macrodipole by introducing a negatively charged residue at S60 and a positively charged residue at D63 would likely have a favorable effect on CD4 thermostability (38), as well as introduce unfavorable electrostatic interactions with the MHC II α chain E88 and K176 (**Figure 1b**). Accordingly, we set out to investigate the effects of alanine, aspartate, glutamate, or lysine substitutions at CD4 positions 60 and 63, hypothesizing that the charged substitutions would have the strongest effects in abolishing MHC II binding.

Normally, CD4 acts together with a TCR to recognize peptide-MHC II complexes (39), but CD4 alone demonstrates low (40) (>2 mM K_D) monomeric affinity for MHC II *in vitro*. Hence, we used a previously reported affinity-matured CD4 (2), CD4_{YW}, to evaluate the effects of our designed CD4 substitutions using surface plasmon resonance (SPR) (41). The Q40Y/T45W substitutions in CD4_{YW} only improve interactions with MHC II β and would not be expected to confound the results of the MHC II α -binding knockdown substitutions. Since CD4 S60R and D63R increased binding to MHC II in the prior CD4 affinity maturation (2), arginine was not included as a potential amino acid substitution at CD4 position 63. With these considerations, we performed surface plasmon resonance (SPR) binding assays using the full ectodomain CD4_{YW} D1-D4 containing various S60/D63 substitutions against MHC II (**Figure 1c**). As hypothesized, the charged substitutions displayed greater reductions in MHC II binding than the alanine substitutions, in rank order of strongest to weakest binding: CD4_{YW} AA > CD4_{YW} AK ~ CD4_{YW} EA > CD4_{YW} EK > CD4_{YW} DK. Remarkably, CD4_{YW} DK's reduced affinity for MHC II appeared to be equivalent to that of the negative control—CD4 that lacks the high-affinity Q40Y/T45W substitutions and therefore has negligible binding to MHC II in SPR assays.

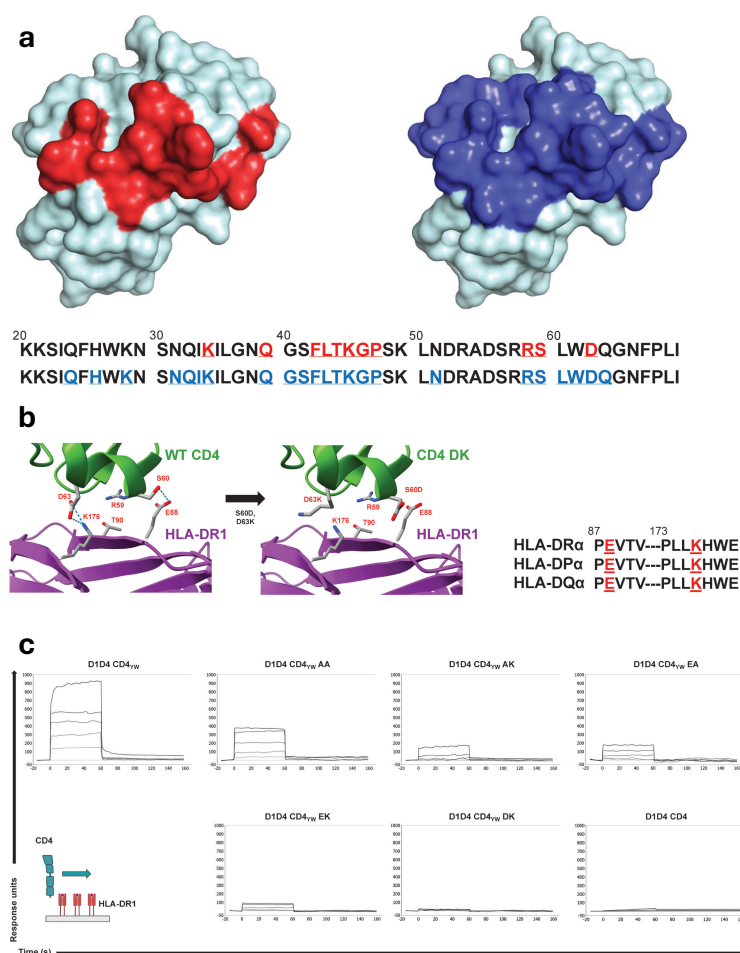


Figure 1. Design of MHC II binding knockdown substitutions for CD4.

(a) Surface representation of the D1 domain of CD4 (residues 1-100) in pale cyan with the MHC II-interacting residues in red on the left and the gp120-interacting residues in blue on the right (PDB: 1GC1). The CD4 sequence is shown in text starting from position 20 and ending at position 70. The MHC II-interacting residues are underlined in red, and below them the gp120-interacting residues are underlined in blue.

(b) Cartoon representation of the interface of CD4 (top; green) and HLA-DRα (bottom; purple) and stick representation of side chains in close proximity (PDB: 3S4S). Dashed lines indicate S60 and D63 of wildtype CD4 (left) interacting with E88 and K176 of HLA-DRα. Modeling the S60D and D63K substitutions in CD4 DK (middle) show unfavorable electrostatic interactions with HLADRα. The sequences of HLA-DRA, HLA-DPA, and HLA-DQA are aligned and shown on the right (aa 87-91 and 173-179), and E88 and K176 are underlined in red.

(c) SPR sensorgrams of HLA-DR1 and variants of the full ectodomain of CD4. Biotinylated HLADR1 was immobilized on the streptavidin-conjugated surface and multi-cycle kinetics was performed using CD4 at various concentrations (100 μ M to 1.23 μ M in 3-fold dilutions). Injections were performed starting with SPR buffer and proceeded from the lowest concentration to the highest concentration. No regeneration was performed between injections. A cartoon of the SPR conditions is shown in the bottom left. CD4 variants with Q40Y and T45W are denoted in subscript, and S60 and D63 substitutions are indicated by the final paired letters (e.g. “CD4_{YW} AA” contains Q40Y, T45W, S60A, and D63A).

3.3.2 Surface and core optimizations improve stability

In parallel, we explored optimizing the surface exposed and buried core amino acids of CD4 to engineer improved thermostability. Analysis of the surface properties of D1D2 was performed in PyMol (36) and the TRIAD physics-based protein design platform to identify design targets for mutagenesis. Our structure-guided analysis identified a hydrophobic patch in D2 that consisted of the side chains L109, L151, V175, and L177 (**Figure 2a**). These residues would normally be partially buried due to contacts with CD4 D3 in the full CD4 ectodomain (42); however, they become fully exposed when CD4 is truncated to L177 as in CD4-Ig. We then performed *in silico* mutational energy score calculations (43, 44) in TRIAD to determine amino acid substitutions that would stabilize this exposed hydrophobic patch (**Extended Data Figure 1**). Nine of the most favorable predictions were expressed and screened for thermostability using differential scanning fluorimetry (DSF) (45) to determine their T_m s. The best performing set of substitutions, L109K, L151K, V175I, and L177D (KKID), improved the T_m of D1D2 by 6°C (**Figure 2b**). We investigated the mechanism for this improvement by modeling L109K, L151K, and L177D, observing that the substituted residues could form a network of salt bridges at the exposed C-terminus of CD4 D1D2 (**Figure 2a**). We also tested a known space-filling A55V substitution (5) in D1 that improved the T_m of D1D2 by 8°C (**Figure 2b**). When the A55V and KKID substitutions were combined, the resulting VKKID variant stabilized D1D2 by 14°C (**Figure 2b**). It is likely that the two sets of substitutions additively improved the T_m by independent mechanisms: the A55V substitution stabilized the hydrophobic core of the D1 domain, and the KKID

substitutions stabilized the hydrophilic surface of the D2 domain. DSF measurement results for all variants are compiled in **Extended Data Table 1**.

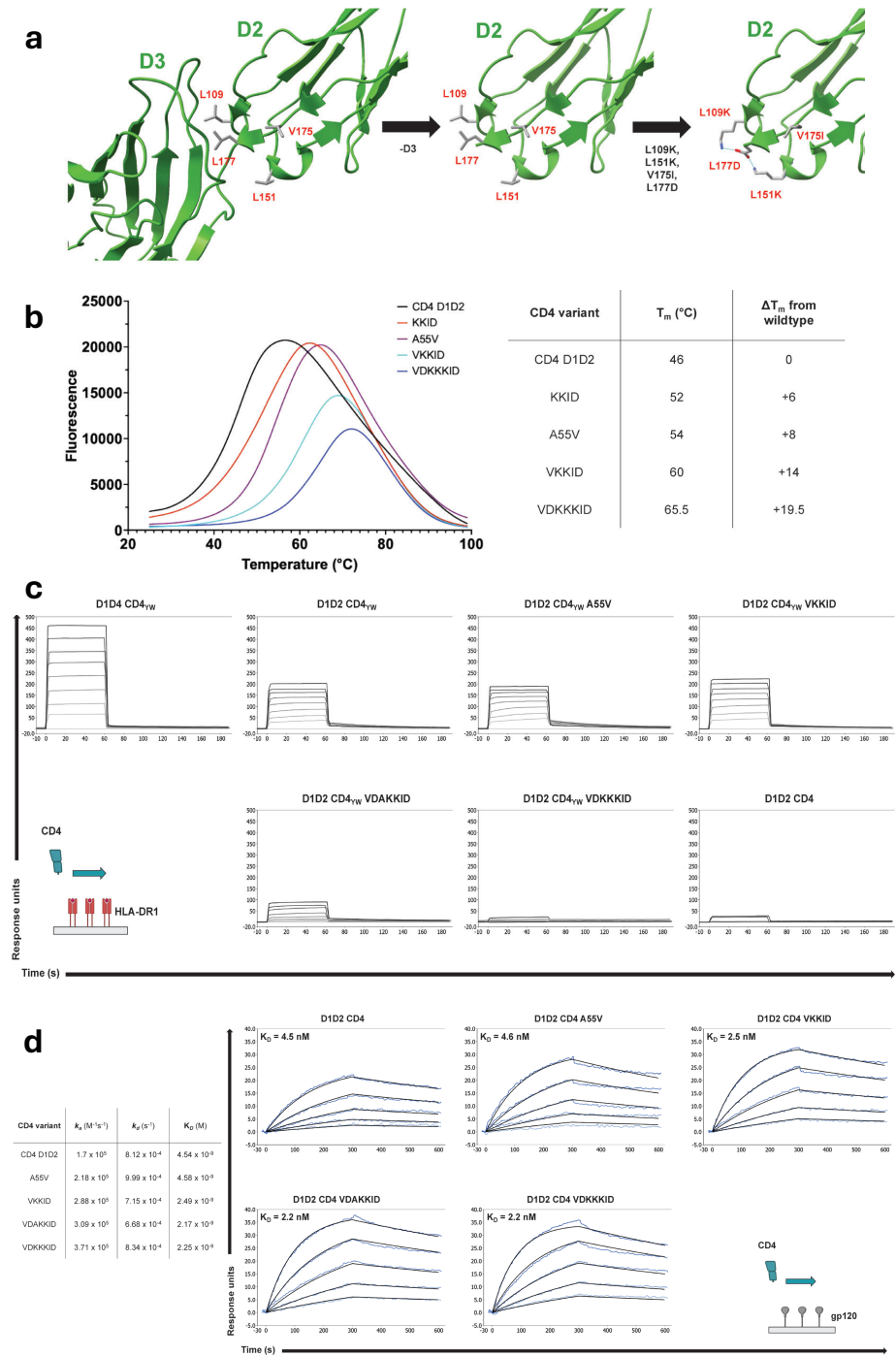


Figure 2. MHC II knockdown and thermostabilizing substitutions preserve CD4's affinity for gp120.

- (a) Cartoon representation of D2 and D3 (left; green) from the full ectodomain of CD4 (PDB: 1WIP) and the side chains of L109, L151, V175, and L177 as sticks. Truncating the D3 domain fully exposes the previously partially buried hydrophobic residues (middle). Modeling the L109K, L151K, V175I, and L177D substitutions (right) shows a newly formed stabilizing network of electrostatic interactions in the KKID variant.
- (b) DSF data of CD4 variants (left) and their melting temperatures in tabular form (right).
- (c), (d) SPR sensorgrams of CD4 variants against HLA-DR1 (c) and gp120 (d).
- (c) Biotylnlated HLA-DR1 was immobilized on streptavidin-conjugated surfaces and CD4 variants were injected for multi-cycle kinetics as in Figure 1c at various concentrations (100 μ M to 0.78 μ M in 2-fold dilutions).
- (d) Full-length BG505 gp120 was immobilized using primary amine coupling and CD4 variants were injected for multi-cycle kinetics at various concentrations (31.25 nM to 1.95 nM in 2-fold dilutions). The gp120 surfaces were regenerated with 10 mM glycine pH 3.0 between injections. Experimental data (blue) were globally fitted using a 1:1 Langmuir binding model (black) and the kinetic rate constants are reported in tabular form (left). Cartoons of the SPR conditions are shown for c (bottom left) and d (bottom right).

3.3.3 gCD4 is thermostable and gp120-specific

After identifying the S60 and D63 substitutions for reducing MHC II binding, we merged these with the VKKID substitutions to make thermostable D1D2 constructs with reduced affinity for MHC II. DSF of the resulting VDKKKID variant indicated that the S60D and D63K substitutions improved thermostability as hypothesized (**Figure 2b**). We then performed SPR with our D1D2 constructs against gp120 and MHC II to confirm that binding to gp120 was retained and only binding to MHC II was knocked down. SPR with MHC II showed that only CD4_{YW} D1D2 variants with substitutions at S60 and D63 had reduced affinity for MHC II, and the results reinforced that S60D and D63K were the most disruptive substitutions (**Figure 2c**). Furthermore, the SPR with gp120 confirmed that all evaluated CD4 D1D2 variants exhibited single-digit nM affinity for gp120, consistent with previously reported values (46, 47) (**Figure 2d**). From these data, we concluded that our targeted substitutions did not disrupt CD4's ability to bind gp120.

Additional mutational screening of KKID revealed that the V175I substitution was unnecessary and that L177E was slightly more thermostable than L177D, perhaps due to glutamate's longer side chain facilitating more favorable salt-bridge formation with the substituted residues in L109K and L151K. We thus defined gp120-specific CD4, or gCD4, as the set of D1D2 substitutions containing A55V, S60D, D63K, L109K, L151K, and L177E (VDKKKE). Compared to wild-type D1D2, gCD4 exhibited minimal evidence of thermal denaturation at physiologic temperatures, an improved T_m (66.5°C) by >20°C, and yielded nearly 50-fold more protein from mammalian cell expressions (**Extended Data Table 1 and Extended Data Figure 2**).

3.3.4 gCD4-Ig inherited the properties of gCD4

Having developed gCD4 with increased thermostability and diminished binding to MHC II, we asked whether these qualities would be inherited if gCD4 were introduced into the CD4-Ig and eCD4-Ig Fc-fusion molecules. DSF of the CD4-Ig and eCD4-Ig variants indicated that gCD4-Ig (T_m =64.5°C) and e-gCD4-Ig (T_m =62.5°C) were 10°C and 9°C more thermostable than CD4-Ig and eCD4-Ig, respectively (**Extended Data Figure 3**). Surprisingly, gCD4-Ig and e-gCD4-Ig had lower melting temperatures than their soluble gCD4 counterpart, likely because the Fc domains undergo thermal denaturation before the gCD4 domains: human Fc initiates its first melting transition around 60-65°C (48), consistent with our observations. We next repeated SPR with MHC II using CD4-Ig variants to investigate how avidity would influence the binding interactions. Similar to the behavior of monomeric CD4, introducing the S60D and D63K substitutions to CD4_{YW}-Ig in the form of gCD4_{YW}-Ig greatly reduced binding to MHC II (**Extended Data Figure 4**). After establishing that gCD4-Ig showed diminished MHC II binding by SPR, we investigated how our CD4-Ig variants would behave in more biologically relevant contexts.

We characterized CD4-Ig's ability to engage MHC II *ex vivo* through flow cytometry of primary human tonsillar leukocytes (49) to account for the surfaces of APCs being densely coated with MHC II (50). We visualized CD4-Ig by introducing an Avitag (51) to the C-

terminus of the Fc, enabling biotinylation and detection with fluorophore-conjugated streptavidin (**Figure 3a**). Clinical trials of PRO-542, a CD4-IgG2 Fc-fusion molecule, achieved serum concentrations of $\geq 1 \mu\text{M}$ with a single 10 mg/kg dose and reported rapid clearance of PRO-542 (11, 12). Mimicking those conditions, tonsil cells were labeled using 1 μM (100 $\mu\text{g/mL}$) of the CD4-Ig variants (**Figures 3b-g**). Flow cytometry with CD4-Ig variants containing the affinity-enhancing Q40Y/T45W substitutions indicated that CD4_{YW}-Ig labeled the majority of CD19⁺ tonsillar B lymphocytes (**Figure 3d**). Comparatively, gCD4_{YW}Ig labeled a smaller proportion of tonsillar B lymphocytes (**Figure 3e**). Without the affinity enhancing substitutions, CD4-Ig also labeled a small proportion of B lymphocytes (**Figure 3f**), an unexpected result due to wild-type CD4's weak monomeric affinity for MHC II (40). Notably, gCD4-Ig weakly labeled tonsillar B lymphocytes (**Figure 3g**), and the measured fluorescence was only slightly above background levels seen with the negative control condition (nonbiotinylated Avitagged CD4_{YW}-Ig) (**Figure 3c**). We observed an identical staining pattern using tonsil cells from an unrelated donor (**Extended Data Figure 5**), supporting the generalizability of the S60D and D63K substitutions in abolishing MHC II binding irrespective of MHC haplotype.

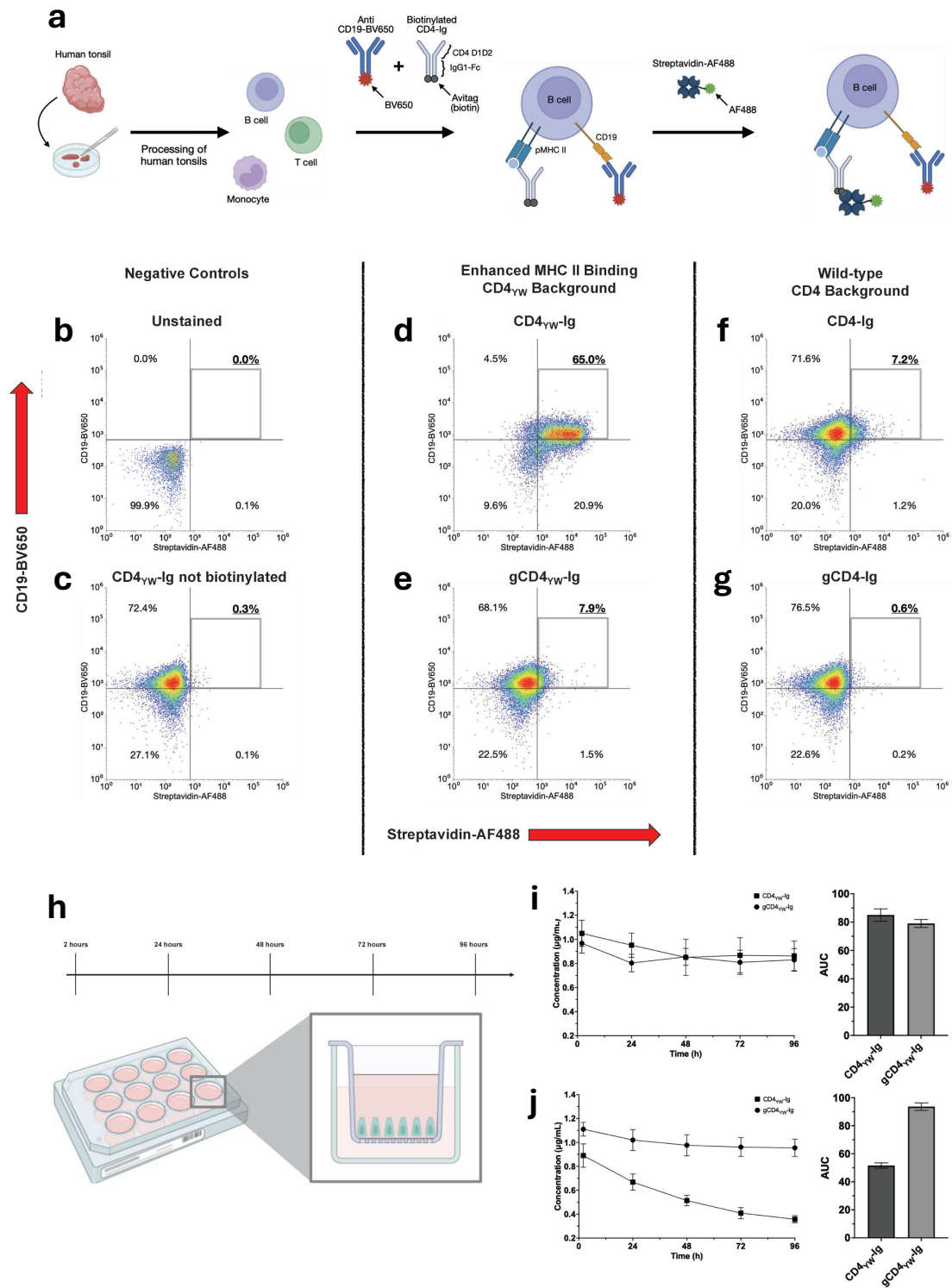


Figure 3. gCD4-Ig persists in tonsil organoid cultures due to reduced MHC II binding.

(a) Diagram of the flow cytometry experimental set-up. Mononuclear leukocytes were isolated by Lymphoprep gradient density separation, aliquoted, and stored frozen. Aliquots were thawed as needed and samples were prepared with live-dead staining and Human TruStain FcX Fc receptor blocking solution. The cells were then labeled with antiCD19 conjugated to Brilliant Violet 650 and biotinylated CD4-Ig variants. After removing excess antibody, the cells were labeled with streptavidin conjugated to Alexa Fluor 488 and analyzed by flow cytometry.

(b)-(g) Flow cytometry plots of tonsillar cells. CD19 signal is plotted on the y-axis and streptavidin signal is plotted on the x-axis. The plots were gated on viable singlet cells, and the fluorescence thresholds were determined using the negative controls (unstained cells and stained cells labeled with unbiotinylated Avitagged CD4_{YW}-Ig). The proportion of double-positive cells for CD19 and CD4-Ig are indicated in the top right quadrant.

(h) Graphic of the tonsil organoid culture experimental set-up. Five to six million tonsil cells were aliquoted in 12-well transwell plates with each well containing a 1 mL outer reservoir of media. 20 μ L from the outer reservoirs were sampled at 2 hours, 24 hours, 48 hours, 72 hours, and 96 hours post-transwell seeding. The outer reservoirs were mixed by pipetting prior to each collection and the timepoints were stored at -80°C.

(i)-(j) Time-course measurements (left) of the concentration of CD4_{YW}-Ig and gCD4_{YW}-Ig incubated in tonsil culture media without tonsil cell seeding (i) or with tonsil cell seeding (j) and their area under the curves (right). Concentrations were determined by ELISA in technical triplicates using the respective proteins as standards. Each point is an average of three independent transwell samples and the error bars are s.d. Variants containing the Q40Y and T45W substitutions are denoted in subscript. CD4_{YW}-Ig for i and j additionally contain the thermostabilizing A55V, L109K, L151K, and L177E substitutions. gCD4 constructs for e, g, i, and j contain A55V, S60D, D63K, L109K, L151K, and L177E substitutions. Donor #664 was used for the data in b-g and donor #669 was used for the data in i and j.

After observing that CD4-Ig measurably binds to cell surface MHC II, we tested our hypothesis that target-mediated drug disposition plays a role in CD4-Ig pharmacokinetics by employing a human tonsil organoid model (49) (**Figures 3h-j**). We estimated that a

leukocyte:CD4-Ig mass ratio of 1000:1 would simulate a typical 15 mg/kg CD4-Ig infusion dose for a 70 kg reference adult male (52), which translates to a concentration of 10 nM (1 μ g/mL) in the tonsil organoids (**Extended Data Figure 6**). To compensate for the 100-fold lower concentration compared to our flow cytometry experiments, we used CD4_{YW}-Ig constructs at 10 nM (**Extended Data Figures 7d and 7e**) as a first approximation for the behavior of CD4-Ig at 1 μ M (**Extended Data Figures 7b and 7c**). Tonsil organoids from a single donor were cultured for four days in the presence of B-cell activating factor (BAFF), which promotes B lymphocyte survival and surface MHC II expression (49, 53). When CD4_{YW}-Ig proteins with and without the MHC II-abrogating DK substitutions were incubated in only culture media, there was no difference in protein loss over time (**Figure 3i**). However, in the presence of tonsil cells, the construct with intact MHC II binding underwent accelerated clearance from the culture, whereas the construct with abrogated MHC II binding did not (**Figure 3j**). These data suggest that MHC II binding contributes to accelerated CD4-Ig turnover in the tonsil organoids.

After characterizing gCD4-Ig's behavior with MHC II-presenting tonsil cells, we evaluated neutralization of a global representative 12-strain panel of HIV-1 pseudoviruses (54) by gCD4-Ig. Although gCD4 did not demonstrate improved binding affinity to gp120 by SPR, its inclusion in CD4-Ig and eCD4-Ig improved the half-maximal inhibitory concentrations (IC₅₀s) of the precursor molecules: gCD4-Ig (geometric mean IC₅₀ 0.74 μ g/mL) and e-gCD4-Ig (geometric mean IC₅₀ 0.035 μ g/mL) were 5.6 and 7.7-fold more potent than CD4-Ig and eCD4-Ig, respectively (**Figure 4a**). CD4-Ig and gCD4-Ig were both not polyreactive in a baculovirus polyreactivity assay (55) (**Extended Data Figure 8**), suggesting that non-specific interactions did not account for the improved neutralization. Instead, we hypothesize that the potency increases may be due to improvements in the conformational stability of gCD4 rather than higher affinity for gp120.

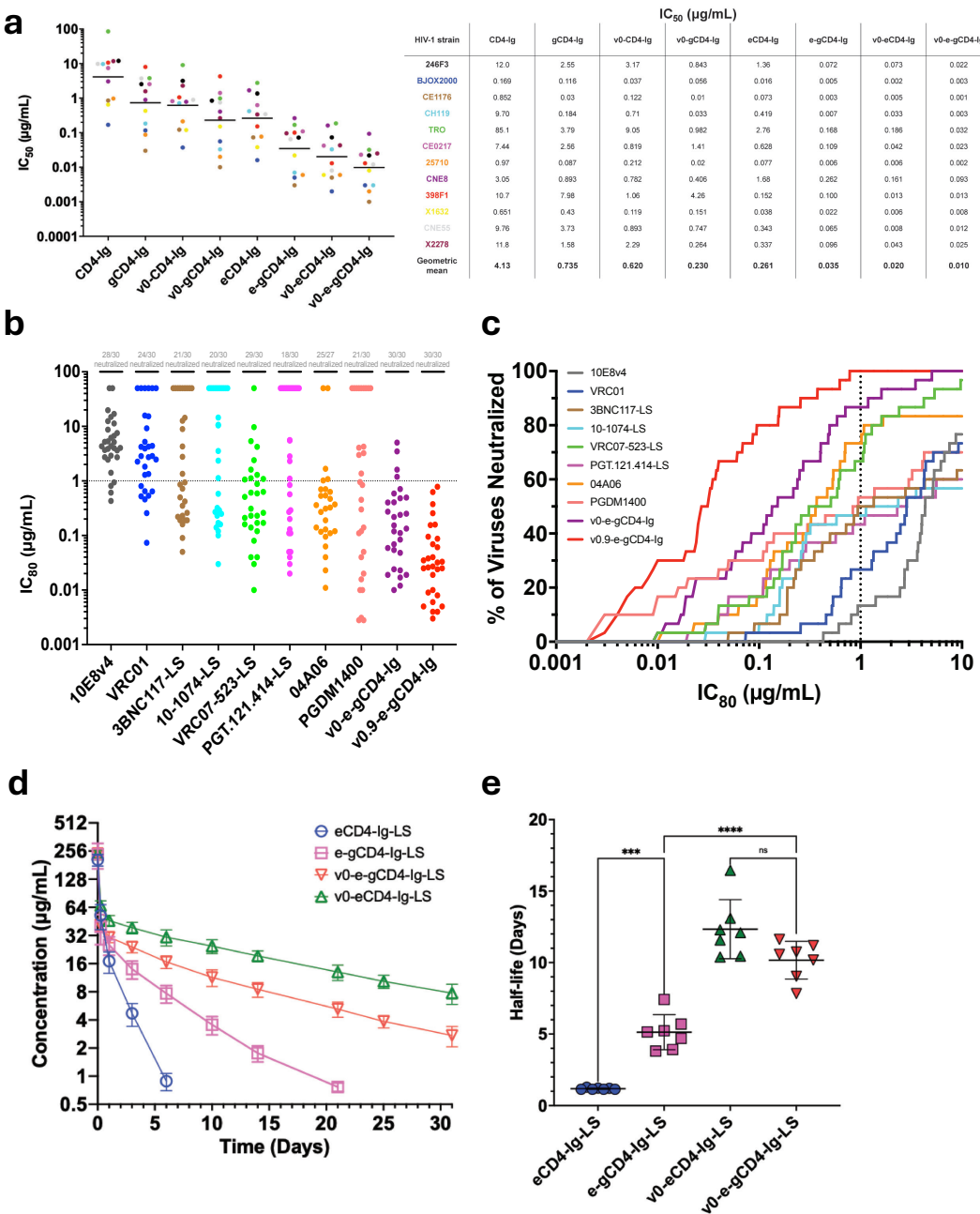


Figure 4. CD4-Ig incorporating gCD4 display enhanced HIV-1 neutralization and have longer half-lives in humanized mice.

(a) IC₅₀ data of CD4-Ig variants against the global 12-strain panel in Tzm-bl cell assays as a scatter plot (left) and table (right). Each point represents the average of technical duplicates for a given virus strain and the black bars indicate the geometric means.

(b) IC₈₀ data of HIV-1 bNAbs and e-gCD4-Ig variants against a panel of 30 viruses from the AMP trials. Non-neutralized viruses are shown as points at 50 µg/mL.

(c) Coverage curves of the viruses from the AMP trials using the data in (b). Each curve indicates the fraction of the virus panel that is 80% neutralized at a given antibody concentration. The protective IC₈₀ threshold from the AMP trials is signified by the dotted lines in (b) and (c) at 1 µg/mL. Neutralization data of the bNAbs were obtained from the CATNAP database.

(d) PK data of eCD4-Ig-LS variants in SCID hFcRn transgenic mice. Each group contained seven mice and the averages of the plasma concentrations are plotted.

(e) Scatter plot of the terminal half-lives of eCD4-Ig-LS variants. Half-lives were determined from individual mice and the average half-lives are presented as black bars. Statistical significance was determined by one-way ANOVA using Tukey's multiple comparisons (***: $p < 0.001$; ****: $p < 0.0001$; ns: not significant). Error bars in (d) and (e) are s.d.

3.3.5 gCD4-Ig variants have therapeutic potential

During these latest experiments, a v0-CD4-Ig containing the CD4 substitutions G6A, S23N, A55V, V128L, R134G, N164D, K167T, and V168L was demonstrated to be more thermostable than wild-type CD4-Ig (14). Since gCD4 only has one substitution in common—A55V—we tested whether the two sets of substitutions were compatible with each other. We first investigated the determinants of v0-CD4-Ig's thermostability by expressing its CD4 D1D2 domains and performing DSF. v0-CD4 displayed a T_m of 78°C, with most of its thermostability contributions coming from the hydrophobic core substitutions G6A, A55V, V128L, and V167L (v0_{core}, T_m of 72°C). The combination of the v0-CD4 and gCD4 substitutions, v0-gCD4, had the highest T_m of 80°C, and we again observed that Fc-fusion constructs had melting temperatures that were lower than the most stable CD4 domains (**Extended Data Table 1**), supporting our hypothesis that the Fc's thermal denaturation occurs before the stabilized CD4 domains. We next introduced the Q40Y/T45W substitutions into v0-CD4 and tested its binding to MHC II by SPR. Similar to the previous SPR results, v0-CD4_{YW} bound to MHC II, and the introduction of the S60D and D63K substitutions in v0-CD4_{YW} DK and v0-gCD4_{YW} significantly reduced MHC II binding

(**Extended Data Figure 9**). We obtained equivalent results with v0.9-CD4_{YW} and v0.9-gCD4_{YW}, which contained additional R59K/K90R substitutions that enhance affinity for gp120¹⁴ (**Extended Data Figure 9**). Taken together, these data suggest that S60D and D63K are necessary and sufficient to reduce binding to MHC II, independent of the extensive amino acid substitutions in v0-CD4.

Since v0-gCD4 combined the favorable properties of v0-CD4 and gCD4, we evaluated whether v0-gCD4 would similarly improve CD4-Ig's neutralization potency. CD4-Ig and eCD4-Ig molecules containing v0-gCD4 demonstrated the highest potency against the 12-strain panel compared to those containing gCD4 or v0-CD4 (**Figure 4a**). The Antibody Mediated Prevention (AMP) trials (16) had determined that protective neutralization titers correlate to *in vitro* IC₈₀ values less than 1 µg/mL; thus, we asked if our most potent variants could neutralize below this threshold by evaluating v0-e-gCD4-Ig and v0.9-e-gCD4-Ig against a panel of 30 AMP trial pseudoviruses. Compared to known highly potent bNAbs, all of which failed to provide complete coverage against the panel, both v0-e-gCD4-Ig and v0.9-e-gCD4-Ig potently neutralized all viruses. The additional gp120 affinity-enhancing R59K and K90R substitutions in v0.9-e-gCD4-Ig improved potency over v0-e-gCD4-Ig as previously reported (14), and v0.9-e-gCD4-Ig neutralized all viruses below the 1 µg/mL threshold (**Figures 4b and 4c and Extended Data Table 2**). From these results, v0.9-e-gCD4-Ig demonstrated superior breadth and potency against AMP viruses than current therapeutic bNAbs, suggesting that passive immunization with v0.9-e-gCD4Ig could provide protection against circulating HIV-1 strains.

Finally, we asked how the enhanced thermostability of v0-gCD4 affected the pharmacokinetics of eCD4-Ig by performing pharmacokinetic studies of eCD4-Ig variants containing the half-life enhancing LS mutations (56) using SCID human FcRn transgenic mice (57, 58). These mice are deficient in MHC II due to impaired lymphocyte development (58) and low levels of APC activation (59), enabling us to minimize MHC II binding as a confounding variable in our half-life estimations. Our studies revealed that eCD4-Ig-LS had a short half-life of ~1 day despite containing the LS mutations, but e-gCD4-Ig-LS and v0-e-

gCD4-Ig-LS demonstrated improved terminal half-lives of ~5 and ~10 days (**Figures 4c and Figure 4d**), respectively. v0-eCD4-Ig had the longest terminal half-life of ~12 days, but this difference was not statistically significant compared to v0-e-gCD4-Ig ($p=0.19$), and both molecules achieved half-lives in mice on par with known therapeutic bNAbs (60). The net charge of an antibody is known to affect its pharmacokinetics (61), although electrostatics may play a minor role in our results as e-gCD4-Ig has a longer half-life than eCD4-Ig despite having more positively charged surface residues. We instead speculate that the relative rank order of the eCD4-Ig variants' half-lives in the SCID mice may be most reflective of their respective CD4 domains' thermostabilities. From these findings, we conclude that thermostabilizing the CD4 domains improved the pharmacokinetics of eCD4-Ig in SCID human FcRn transgenic mice.

3.4 Discussion

Here, we provide evidence that CD4's affinity for MHC II can be abolished while retaining CD4's affinity for HIV-1, paving the way for the clinical development of CD4-containing biologics incorporating gCD4. Using CD4-Ig as a proof-of-principle, our data demonstrated that we have produced HIV-1 entry inhibitors with pharmacokinetic properties and neutralizing capabilities that surpass those of therapeutic bNAbs. Encouragingly, our most potent variant neutralized all tested pseudoviruses from the AMP trials at titers indicative of protective efficacy, a feat that would normally require a mixture of different bNAbs (17). Other CD4-containing biologics, such as bNAbs-CD4 fusions (8), CD4 nanoparticles (20), and multimeric CD4-Fc constructs (5) are likely to benefit from the inclusion of gCD4 in their future constructions. Finally, our computational and rational design approaches can be reasonably applied to other proteins to engineer their biophysical and biochemical properties.

Our observation that CD4-Ig binds to cellular MHC II raises several considerations for the clinical use of CD4-containing biologics as it remains unknown how such molecules might interact with human physiology. A previous study indicated that soluble monomeric CD4 did not acutely inhibit CD4⁺ T cell activation *in vitro* (62), but it is possible that multimeric molecules such as CD4-Ig would compete for MHC II engagement and interfere with CD4⁺

T cell activation and/or development due to greater avidity. It is also not known if CD4-Ig would indiscriminately mediate Fc receptor-dependent cytotoxicity against MHC II-containing cells. These unknowns motivated precaution in the clinical formulation of PRO-542, which contained four copies of CD4 fused to IgG2 Fc to weaken Fc effector functions (6). However, these concerns are now mitigated by gCD4's near-complete specificity for HIV-1, which should enable safer implementation of CD4-Fc fusions with intact effector functions. Fc effector functions are key features of protective antibodies, allowing them to kill virally infected cells through mechanisms such as antibody-dependent cellular cytotoxicity (ADCC). For example, it was shown that eCD4Ig with IgG1, but not IgG2, Fc can mediate ADCC of HIV-1-infected cells *in vitro* (63). Fc effector functions are also implicated in the ability of bNAbs to prolong control of viremia via the vaccinal effect (64). Due to these advantages, biologics with IgG1 Fc would be preferred for developing curative therapeutic regimens for people living with HIV-1. In addition to treating active infection, it has been shown that eCD4-Ig delivered by AAV can act like an HIV-1 vaccine and provide years-long protection from HIV-1 infection in infant rhesus macaques (65). In the context of AAV gene therapy, reducing MHC II binding may become a more salient consideration for potentially lifelong eCD4-Ig expression.

3.4.1 Limitations of the study

Limitations of our study include the functional relevance of our tonsil and mouse experiments in the context of human physiology and the translatability of our *in vitro* neutralization results to *in vivo* efficacy. First, our flow cytometry experiments likely underestimate how much CD4-Ig binds to MHC II *in vivo*. The rapid dissociation kinetics of CD4 and MHC II as seen in our SPR data imply that a large proportion of CD4-Ig would readily disengage from the tonsillar B cells during washing and that the binding at equilibrium would be higher than what we have recorded. Second, the effect size of MHC II-mediated clearance remains unknown physiologically, although our tonsil organoid data suggest that MHC II binding results in measurably faster clearance. Transgenic mice expressing fully human MHC II may be an appropriate model to test this variable. Third, the preparation of tonsil organoids results in a limited composition of mononuclear white blood

cells that are predominantly B and T lymphocytes. Macrophages express MHC II and make up a much larger proportion of the human immune system by mass than any other individual white blood cell (52). Therefore, macrophages may significantly contribute to the clearance of CD4-Ig *in vivo* through the reticuloendothelial system. Since spleens contain large numbers of macrophages, switching to spleen-derived organoids may address this limitation (49, 52). Fourth, we did not test whether CD4-Ig binding to MHC II has functional relevance, such as investigating B and T lymphocyte activation in the tonsil organoids longitudinally. Finally, although our e-gCD4-Ig variants displayed excellent neutralization breadth and potencies, we did not evaluate if they prevent or treat HIV-1 infection *in vivo*. However, the comparatively less potent eCD4-Ig was shown to prevent HIV-1 infection in mice and rhesus macaques (7, 13, 65); thus, it is likely that our molecules will behave similarly although further studies are required to investigate the efficacies of e-gCD4-Ig variants. While more challenges remain, gCD4 demonstrates great potential to develop superior CD4-containing biologics to provide safe, durable, and broad protection and/or treatment against HIV-1.

3.5 Materials and Methods

3.5.1 DNA constructs and vectors

Human codon-optimized CD4 sequences with a GS linker and C-terminal 6xHis tag were designed and ordered as gBlocks from IDT using IDT's codon optimization tool. The gBlocks were cloned into the vectors pTT5 (National Research Council of Canada) or p3BNC through Gibson assembly (66) using NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs) according to manufacturer recommendations. DNAs were transformed into chemically competent Mix and Go! DH5α *E. coli* (Zymo Research) and grown on LB agar plates containing 100 µg/mL carbenicillin. Plasmids were isolated from *E. coli* using Zippy Plasmid Miniprep Kits (Zymo Research) and verified through Oxford nanopore sequencing (Plasmidsaurus).

CD4-Ig and eCD4-Ig constructs were designed as previously reported. gBlocks were ordered from IDT and assembled in p3BNC as above. All CD4-Ig constructs contained identical mouse VH leader sequences. The sequence for human TPST2 was obtained from Uniprot

(Accession ID: O60704) and cloned into p3BNC with its native Golgi signal sequence. For Avitagged proteins, the 15 amino acid Avitag sequence GLNDIFEAQKIEWHE was added to the C-terminus of the Fc. The sequence for *E. coli* biotin ligase was assembled in p3BNC as above and contained a mouse VH leader sequence and a C-terminal KDEL endoplasmic reticulum retention motif. Plasmid DNAs for mammalian cell transfection were prepared from *E. coli* using Macharey-Nagel midiprep and maxiprep kits.

The pTT5 plasmid containing His-tagged full-length BG505 gp120 was obtained from the Bjorkman lab.

3.5.2 Protein expression and purification

Gibco Expi293F cells were transfected with 1 µg of DNA/mL of cells using the Gibco Expi-Fectamine transfection kit according to manufacturer recommendations. Four days post-transfection, cultures were centrifuged and supernatants were filtered through 0.45µm filters. For His-tagged proteins, Pierce disposable columns were packed with HisPur Ni-NTA resin (Thermo Scientific), rinsed with water, and equilibrated with Gibco 1X Dulbecco's PBS with no magnesium and no calcium, pH 7.4 (DPBS). Clarified supernatant was applied to equilibrated Ni-NTA resin under gravity flow. Columns were washed with ten column volumes of wash buffer (DPBS, 15 mM imidazole) and eluted with two to three column volumes of elution buffer (DPBS, 300 mM imidazole). Eluted proteins were concentrated using Amicon spin concentrators and further purified via FPLC using a Superdex 200 Increase 10/300 GL column equilibrated with DPBS at 4°C. Fractions containing the desired protein were pooled, concentrated, filtered with a 0.22 µm filter, and stored at 4°C. Protein concentrations were determined by measuring absorbance at 280 nm on a NanoDrop (Thermo Fisher Scientific). Extinction coefficients were calculated by inputting the amino acid sequences of the mature proteins in ExPASy ProtParam.

Expression and purification of CD4-Ig and eCD4-Ig was performed as previously described (14). Plasmids containing human TPST2 and eCD4-Ig were co-transfected at a 1:4 DNA mass ratio of TPST2:eCD4-Ig. Expression of biotinylated CD4-Ig was performed in an

identical manner using a 1:4 ratio of BirA:Avitagged CD4-Ig. Purification of Fc-containing proteins was performed using HiTrap MabSelect SuRe columns (Cytiva) connected to a peristaltic pump according to manufacturer recommendations. Columns were washed with ten column volumes of 1X DPBS and eluted using IgG elution buffer (Pierce) into 1 M Tris-HCl buffer, pH 9 (NC9904344; Boston Bioproducts). Eluted proteins were further purified via FPLC as above.

3.5.3 Computational design of CD4

For *in silico* predictions of thermostabilizing substitutions, the coordinates of CD4 D1D2 were extracted from cocrystal structures of CD4 D1D2 bound to gp120 using PyMOL (PDBs: 1GC1, 2NY1, 2NY2) (1, 67). Water molecules were not included. Rosetta (43) and PHOENIX (44) force fields were used in TRIAD to separately standardize and analyze the extracted CD4 structures. Monte Carlo optimization with sampling was performed without covalent terms on fixed backbones from initial rotamers using the expanded Dunbrack rotamer library (68) and floating nearby residues. Sequence logos of the predicted sequences were generated using WebLogo (<https://weblogo.berkeley.edu/>). A subset of *in silico* predictions were selected for experimental testing based on the frequency of each residue at each position as well as the relative rank order of the unique combinations.

3.5.4 Differential scanning fluorimetry

DSF assays were performed by preparing triplicate samples in DPBS to a final concentration of 5 to 10 μ M of protein and 20X Sypro Orange dye (Invitrogen). A CFX96 Touch Real-Time PCR Detection System (Bio-Rad) was used to measure the fluorescence of the samples by heating from 25.0°C to 99.0°C in 0.5°C increments and 30 second equilibration times. Background subtraction was performed using the mean fluorescence of triplicate samples containing only buffer and dye. The data were exported and plotted in GraphPad Prism 11. Apparent melting temperatures were defined as the inflection point of the first derivative of the fluorescence data with respect to time.

3.5.5 Surface Plasmon Resonance

Analyte proteins for SPR injection were prepared as described above with one modification: the FPLC buffer was substituted for 1X HBS-EP+ (150 mM NaCl, 10m HEPES pH 7.4, 2mM EDTA, 0.05% Tween-20) prepared from 20X HBS-EP+ buffer (Teknova). Multi-cycle kinetics were performed at 25°C in a Bruker Sierra SPR Pro-32 instrument. SPR running buffer was 1X HBS-EP+ and SPR wash buffer consisted of 0.05% Tween-20. The ligand for SPR experiments in **Figures 1c** and **2c** was biotinylated monomeric HLA-DR1 loaded with influenza A hemagglutinin peptide 306-318 (PKYVKQNTLKLAT), obtained from the NIH Tetramer Core Facility. The ligand for subsequent SPR experiments was biotinylated monomeric HLA-DR1 loaded with SARSCoV-2 S peptide 873-888 (GAALQIPFAMQMAYRFF), also obtained from the NIH Tetramer Core Facility. Biotinylated ligands were diluted to 100 nM in running buffer and directly immobilized on a Bruker Biotin-Tag Capture Sensor to achieve response units (RU) of approximately 500-1,000. No additional protein was immobilized on reference flow cells as they already contained streptavidin. Three start-up injections of buffer preceded analyte injections. Analyte samples were prepared in 2-fold or 3-fold dilution series in running buffer and injected at a flow rate of 30 uL/min followed by a disassociation phase. CD4 and CD4-Ig exhibited extremely rapid dissociation kinetics and required no surface regeneration between cycles. Analyte injections were performed in the order from lowest concentration first to highest concentration last. For SPR of CD4 and gp120, full-length BG505 gp120 was diluted to 100 nM in 10 mM sodium acetate pH 4.5 and immobilized on a Bruker high-capacity amine chip to approximately 400-900 RU according to manufacturer recommendations. M8a-3 IgG, an anti-SARS-CoV-2 antibody, was equivalently coupled to reference flow cells (69). Injections were performed with similar conditions as above and included a regeneration step with 10 mM glycine pH 3.0. SPR sensorgrams for all experiments were double subtracted from the reference and buffer injections. Experimental data for SPRs with gp120 were globally fit to a 1:1 Langmuir model using the Bruker Sierra Analysis software.

3.5.6 Baculovirus assay

Polyreactivity was assessed using an ELISA-based method to assay non-specific binding to baculovirus particles as described (55, 70). Briefly, a 1% solution of baculovirus particles in 100 mM sodium bicarbonate buffer, pH 9.6 was adsorbed to a 384-well Maxisorp ELISA plate overnight at 4°C. After blocking with 0.5% bovine serum albumin (BSA) in PBS, CD4-Ig samples and controls were added to the blocked assay plate and incubated for 3 hours at room temperature. The plate was washed with PBS and an HRP-conjugated anti-IgG secondary antibody (Southern Biotech 2015-05) was added and incubated for one hour at room temperature. Plates were washed with PBS and bound IgG was detected by measuring luminescence at 425 nm after addition of SuperSignal ELISA Femto Maximum Sensitivity Substrate (ThermoFisher).

3.5.7 Pseudovirus Neutralization

HIV-1 pseudovirus neutralization assays were performed and analyzed as described (71, 72), either in-house (73) or through the Collaboration for AIDS Vaccine Discovery (CAVD) core neutralization facility. Briefly, HIV-1 pseudoviruses were generated by cotransfecting HEK 293T cells with an HIV-1 Env expression plasmid and a replication-defective backbone plasmid (71, 72). Supernatants were harvested at 72 hours, filtered through a 0.45 μ m sterile filter, then titered using TZM-bl target cells, which express HIV-1 receptors CD4 and CXCR5 (71). CD4-based inhibitors were evaluated in duplicate with a 5-fold, 8-point dilution series of the inhibitor. The CD4-based inhibitor was incubated with pseudovirus for one hour at 37°C then added to TZM-bl target cells. After 48 hours at 37°C, cells were lysed with BriteLite luminescence substrate and luminescence was measured to evaluate the amount of infection at each concentration of inhibitor. The percent neutralization was calculated and plotted as a function of concentration. IC₅₀ and IC₈₀ values, the concentrations at which half-maximal or 80% neutralization were observed, were derived using a nonlinear regression analysis using the Antibody Database (74). These values were exported and plotted in GraphPad Prism 11.

3.5.8 Tonsil Organoids

De-identified human tonsils were obtained from the Cooperative Human Tissue Network and processed as previously reported (49). Briefly, tonsils were sectioned and manually homogenized into suspension. Mononuclear leukocytes were isolated by Lymphoprep density gradient separation, washed with complete media (RPMI with glutamax, 10% Fetal Bovine Serum [FBS], 1X nonessential amino acids, 1X sodium pyruvate, 1X penicillin–streptomycin, 1X Normocin [InvivoGen] and 1X insulin/selenium/transferrin cocktail [Gibco]), and frozen in 1 mL aliquots of 50–100 million cells in FBS + 10% dimethyl sulfoxide (DMSO). Frozen aliquots were stored at -140°C. Tonsil cells were prepared for culturing by thawing in a 37°C water bath and resuspending in 10 mL of prewarmed thaw media. Resuspended cells were centrifuged at 500 x g for 10 minutes at 4°C and the thaw media was aspirated. Pelleted cells were resuspended in prewarmed complete media, and the number of viable cells was determined with Trypan blue using a hemocytometer. The cell suspension was then adjusted to a density of ~60 million cells/mL using complete media. Culture plates were prepared by adding 1 mL of prewarmed complete media supplemented with 1 µg/mL of human B cell-activating factor (BAFF; BioLegend) and, if applicable, CD4-Ig proteins to 12-well plates (Corning). 12 mm diameter Millicell culture plate inserts (MilliporeSigma) were placed into each well and 100 µl of cell suspension was plated into the inserts. For media-only wells, 100 µl of complete media was added to the inserts. Cultures were incubated at 37°C and 5% CO₂ with humidity. Aliquots of culture media were taken from the exterior reservoir after 2 hours, 24 hours, 48 hours, 72 hours, and 96 hours and frozen at -80°C for later analysis. The remaining volumes of blank wells filled with complete media were tracked using a micropipette to estimate evaporation over time in the cultures and cross-verified with the experimental wells at the end of the experiments. ELISA signals were multiplied with a correction factor for evaporation at each time point, accounting for approximately 50 µL of evaporation every 24 hours. Timepoints and the area under the curves were plotted using GraphPad Prism 11.

3.5.9 Flow Cytometry

Frozen tonsil aliquots were thawed in a 37°C water bath and centrifuged for 3 minutes at 1,000 x g at 4°C. The media was discarded, and the cells were washed by resuspending the cell pellet with DPBS. The cells were pelleted and stained with Zombie Violet Fixable Viability Kit (423113; BioLegend) according to manufacturer recommendations in DPBS. After viability staining, cells were washed once with FACS buffer (DPBS, 0.1% Bovine Serum Albumin [BSA; 50-178-1669; Sigma Aldrich]) and cellular Fc receptor was blocked with Human TruStain FcX (422302; BioLegend) according to manufacturer recommendations in FACS buffer. Approximately 1-2 million cells were aliquoted into individual samples of 100 µL of FACS buffer and incubated with 3 µL of Brilliant Violet 650 anti-human CD-19 antibody conjugate (302238; BioLegend) and 1 µM purified Avidin-tagged CD4-Ig proteins for 1 hour on ice protected from light. Samples were pelleted and resuspended in 50 µL of FACS buffer containing streptavidin-Alexa Fluor 488 conjugate (S32354; Invitrogen) diluted 1:50 (final concentration 40 µg/mL). After 30 minutes of incubating on ice protected from light, the samples were washed once with 200 µL of FACS buffer and kept pelleted on ice until the time of measurement. Samples were resuspended in 500 µL of FACS buffer, filtered through 70 µm Flowmi pipette tip strainers (H13680-0070; SP Bel-Art), and immediately analyzed on a Sony SH800S Cell Sorter. Compensation matrices were calculated using unstained and single-stain controls and 50,000 to 100,000 total events were recorded for each sample. Cells were gated for singlets, viability, and the detection of Brilliant Violet 650 and streptavidin-Alexa Fluor 488. Flow cytometry plots were generated using Floreada.

3.5.10 ELISA

Determination of the CD4-Ig concentration in the tonsil organoid cultures was performed by capture ELISA. Clear, flat-bottom, high bind 96-well polystyrene plates (Corning) were coated overnight at 4°C with 50 µL of mouse anti-human CD4 antibody, clone RPA-T4 (300502; BioLegend), diluted to 2 µg/ml in 0.1 M NaHCO₃, pH 9.6. Wells were emptied by dumping into a sink and then blotted on clean paper towels. Coated wells were blocked with 300 µL of blocking buffer (DPBS, 3% BSA, 0.1% Tween-20) at room temperature (RT) for

2 hours. Frozen aliquots of the tonsil culture media samples were thawed and diluted 1:100 into blocking buffer. The blocked plates were then dumped and blotted, followed by application of 50 μ L of the diluted tonsil culture media samples in triplicate. Standard curves of purified CD4-Ig variants were applied on the same plate in triplicate using 4-fold serial dilutions in blocking buffer starting from a top concentration of 1 μ g/mL. The plates were incubated at room temperature (RT) for 2 hours and washed three times with washing buffer (DPBS, 0.1% Tween-20). HRP-conjugated goat anti-human IgG Heavy Chain Superclonal antibody (A56862; Invitrogen) was prepared by diluting 1:10,000 in blocking buffer, and 50 μ L of secondary antibody was applied to each well and incubated at RT for 30 minutes. After three washes, 50 μ L of 1-Step Ultra TMB-ELISA substrate (34028; Thermo Scientific) was applied to each well and incubated at RT for ten minutes protected from light. The reaction was stopped by adding 50 μ L of 1 N HCl to each well and absorbance at 450 nm was measured with a plate reader. The concentration of CD4-Ig in the tonsil samples was estimated by interpolating from standard curves using a four-parameter nonlinear regression model in GraphPad Prism 11. Correction factors from the measured evaporation were applied to the estimated concentrations.

3.5.11 Murine pharmacokinetics

All experimental work relating to mice was performed by The Jackson Laboratory. Four groups of SCID hFcRn mice (JAX strain #018441) at 6-8 weeks of age were intravenously given 10 mg/kg of eCD4-Ig-LS variants. Whole blood was collected at 5 minutes, 6 hours, 1 day, 3 days, 6 days, 10 days, 14 days, 21 days, 25 days, and 31 days from the tail vein and processed to plasma before being stored at -80°C. eCD4-Ig-LS concentrations were assayed by capture ELISA using an anti-CD4 capture antibody and a HRP-conjugated goat anti-human secondary antibody as above. PK analysis was performed using Phoenix WinNonLin 8.5.2.4.

3.6 Acknowledgments

We thank Lisa Wagar and Sam Kim (UC Irvine) for help in establishing tonsil organoid systems, Jost Vielmetter and the Caltech Protein Expression Center for assistance with protein expression and purification, Jost Vielmetter and Harry Gristick for discussions about SPR, Rebecca Hill (Kaiser Permanente School of Medicine) for help with performing power analyses in designing mouse experiments, the NIH Tetramer Core Facility for providing biotinylated HLA-DR1 monomers, and Mike McCune (Gates Foundation) and members of the Mayo and Bjorkman labs for helpful discussions. This work was supported by the Bill and Melinda Gates Foundation INV-057196 (P.J.B.) and INV-036842 (M.S.S.).

3.7 Author Contributions

A.J.B., L.F.C., P.J.B., and S.L.M. conceived the study and P.J.B. and S.L.M. provided funding support. A.J.B., L.F.C., P.J.B., and S.L.M. designed experiments. M.S.S. and P.N.P.G. performed HIV-1 neutralization assays, J.R.K. performed polyreactivity assays. A.J.B. and L.F.C. performed and analyzed all other experiments. A.J.B., L.F.C., P.J.B., and S.L.M. wrote the manuscript.

3.8 Competing Interest Statement

A.J.B., L.F.C., P.J.B. and S.L.M. are inventors on a provisional US patent application filed by the California Institute of Technology that covers the work described in this manuscript.

3.9 Extended Data

Extended Data Table 1. Apparent melting temperatures of CD4-containing constructs.

Samples were measured in technical triplicate and the averages are reported.

CD4 D1D2 Variant	T_m (°C)	ΔT_m from CD4 D1D2 (°C)
CD4 D1D2	46	N/A
L109S L151R V175I L177E (SRIE)	42	-4
L109K L151R L177S (KRVS)	44	-2
L109S L151R V175R L177E (SRRE)	45	-1

L96V	45.5	-0.5
L109K L151K V175T L177S (KKTS)	46	0
L109K L151K V175I L177S (KKIS)	46.5	+0.5
L109Q L151Q V175I L177S (QQIS)	47	+1
L109Y L151E V175T L177H (YETH)	47.5	+1.5
L109Y L151E L177H (YEVH)	47.5	+1.5
L151K L177D (LKVD)	51.5	+5.5
L151R L177D (LRVD)	52	+6
L109K L177 (KLVD)	52	+6
L109K L151K V175I L177D (KKID)	52	+6
L109R L177D (RLVD)	52.5	+6.5
L109K L151R V175I L177D (KRID)	53.5	+7.5
A55V	54	+8
L109K L151K V175I L177E (KKIE)	54	+8
L109R L151K V175I L177D (RKID)	54	+8
L109K L151K L177D (KKVD)	54.5	+8.5
L109K L151K L177E (KKVE)	54.5	+8.5
S23N	55	+9
L109K L151K V175R L177D (KKRD)	55	+9
L109K L151K V175K L177D (KKKD)	55.5	+9.5
I24T	56.5	+10.5
S23N I24T (NT)	58.5	+12.5
S60D D63K L109K L151K V175I L177D (DKKKID)	58.5	+12.5
S60D D63A L109K L151K V175I L177D (DAKKID)	59	+13
S23N A55V L96V (3mut)	59.5	+13.5

A55V L109K L151K V175I L177D (VKKID)	60	+14
A55V S60D D63K L109K C130V L151K C159V L177E (C130V C159V gCD4)	61	+15
A55V S60D D63K (VDK)	62.5	+16.5
S23N A55V S60D D63K L96V (3mut DK)	62.5	+16.5
A55V S60D D63A (VDA)	63.5	+17.5
S23N A55V S60D D63K L96V L109K L151K V175I L177D (3mut DKKKID)	64.5	+18.5
S23N R134G N164D K167T ($v0_{\text{surface}}$)	64.5	+18.5
A55V S60D D63K L109K L151K V175I L177D (VDKKKID)	65.5	+19.5
A55V S60D D63K L109K L151K V175I L177D (VDAKKID)	66	+20
S23N I24T A55V S60D D63K L96V L109K L151K V175I L177D (3mut I24T DKKKID)	66	+20
A55V S60D D63K L109K L151K L177E (VDKKKE/gCD4)	66.5	+20.5
A55V S60D D63A L109K L151K L177E (VDAKKE)	67.5	+21.5
G6A A55V V128L V168L ($v0_{\text{core}}$)	72	+26
S23N S60D D63K L109K R134G L151K N164D K167 L177E ($v0_{\text{surface}}$ gCD4)	72.5	+26.5
G6A A55V S60D D63K L109K V128L L151K V168L L177E ($v0_{\text{core}}$ gCD4)	76.5	+30.5
G6A A55V S60D D63K L109K V128L V168L L177E ($v0_{\text{core}}$ DKKE)	77	+31
G6A A55V S60D D63K L109K V128L R134G L151K V168L L177E ($v0_{\text{core}}$ gCD4 R134G)	78	+32

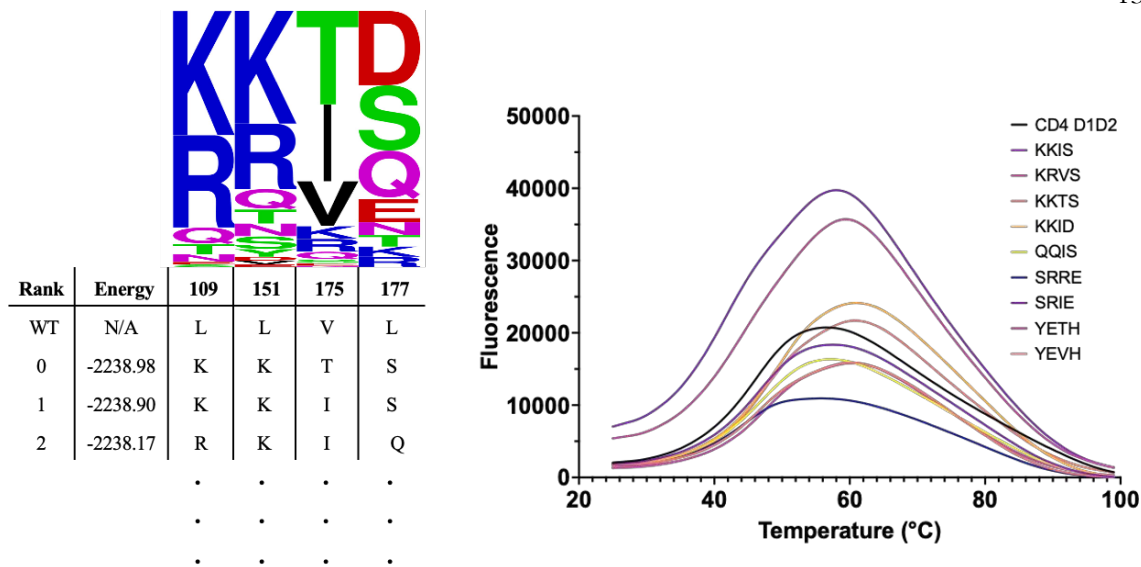
G6A A55V S60D D63K L109K V128L R134G V168L L177E (v0 _{core} DKGKE)	78.5	+32.5
G6A S23N A55V V128L R134G N164D K167T V168L (v0)	78.5	+32.5
G6A S23N A55V L109K V128L R134G L151K N164D K167T V168L L177E (v0 KKE)	78.5	+32.5
G6A S23N A55V S60D D63K R134G N164D K167T V168L L177E (v0 DK)	80	+34
G6A S23N A55V S60D D63K L109K V128L R134G L151K N164D K167T V168L L177E (v0 gCD4)	80	+34
G6A S23N I24T A55V S60D D63K L109K V128L R134G L151K N164D K167T V168L L177E (v0 I24T gCD4)	80	+34
CD4 D1-D4 Variant	T_m (°C)	ΔT_m from CD4 D1-D4 (°C)
CD4 D1-D4	52.5	N/A
S60A D63A (AA)	53	+0.5
S60A D63K (AK)	53	+0.5
S60D D63K (DK)	53.5	+1
S60E D63K (EK)	53.5	+1
S60D D63A (DA)	54	+1.5
S60E D63A (EA)	54	+1.5
CD4-Ig variant	T_m (°C)	ΔT_m from CD4- Ig (°C)
CD4-Ig	54.5	N/A
gCD4-Ig	64.5	+10
v0 CD4-Ig	67	+12.5
v0 gCD4-Ig	67.5	+13

eCD4-Ig variant	T_m (°C)	ΔT_m from eCD4-Ig (°C)
eCD4-Ig	53.5	N/A
e-gCD4-Ig	62.5	+9
v0 eCD4-Ig KKE	65	+11.5
v0 eCD4-Ig DK	65.5	+12
v0 e-gCD4-Ig	65.5	+12

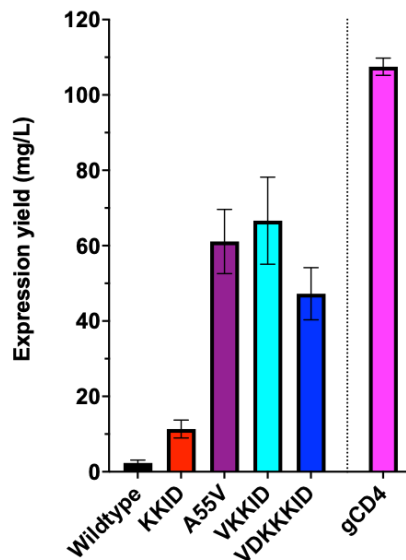
Extended Data Table 2. IC₅₀, IC₈₀, and MPI (maximum percent inhibition) values of v0-egCD4-Ig and v0.9-e-gCD4-Ig against 30 AMP trial pseudoviruses.

	v0-e-gCD4-Ig			v0.9-e-gCD4-Ig		
Virus ID	IC ₅₀ (µg/mL)	IC ₈₀ (µg/mL)	MPI	IC ₅₀ (µg/mL)	IC ₈₀ (µg/mL)	MPI
H703_0013_090Es	0.009	0.019	100	0.001	0.004	100
H703_0203_081_RE_e8A5s	0.040	0.239	100	0.007	0.038	100
H703_0279_110s	0.027	0.111	100	0.007	0.025	100
H703_0472_030s	0.116	0.481	100	0.016	0.061	100
H703_0537_110s_4H1	1.164	5.050	100	0.274	0.782	100
H703_0566_0160s	0.028	0.084	100	0.008	0.027	100
H703_0646_051sN	0.017	0.069	100	0.006	0.026	100
H703_0739_110s	0.005	0.017	100	0.002	0.005	100
H703_0842_200Es	0.122	0.406	100	0.026	0.087	100
H703_0926_070S_2H2	0.259	1.188	100	0.113	0.376	100
H703_1104_100_RE_e10G5s	0.047	0.153	100	0.010	0.033	100
H703_1313_040s	0.007	0.019	100	0.001	0.003	100

H703_1471_190s	0.042	0.134	100	0.011	0.040	100
H703_1675_080s_G613S	0.206	0.709	100	0.043	0.155	100
H703_1750_140Es	0.012	0.048	100	0.002	0.008	100
H703_1764_250_RE_cs	0.019	0.059	100	0.004	0.010	100
H703_1789_230_RE_e3A3s	0.177	0.496	100	0.034	0.095	100
H703_2018_240_RE_e6A1s	0.695	3.479	100	0.136	0.624	100
H703_2117_110_RE_e2A10s	0.066	0.218	100	0.009	0.025	100
H703_2149_060_RE_eB10s	0.056	0.251	100	0.008	0.024	100
H703_2304_150_RE_cs	0.022	0.054	100	0.005	0.019	100
H703_2539_070_RE_e6F6s	0.089	0.273	100	0.017	0.069	100
H703_2631_150_RE_e2F8s	0.003	0.012	100	0.001	0.005	100
H703_2788_030Es_B1	0.007	0.024	100	0.002	0.006	100
H703_2805_080Es	0.040	0.119	100	0.014	0.032	100
V703_0217_050_RE_pplib002_s	0.199	1.608	100	0.065	0.256	100
V703_0629_150_RE_pplib002_s	0.004	0.010	100	0.002	0.004	100
V703_1298_080_RE_pplib002_s	0.121	0.587	100	0.028	0.158	100
V703_1848_190_RE_pplib002_s	0.004	0.023	100	0.002	0.009	100
V703_1915_250_RE_sgaA3_s	0.117	0.403	100	0.013	0.035	100

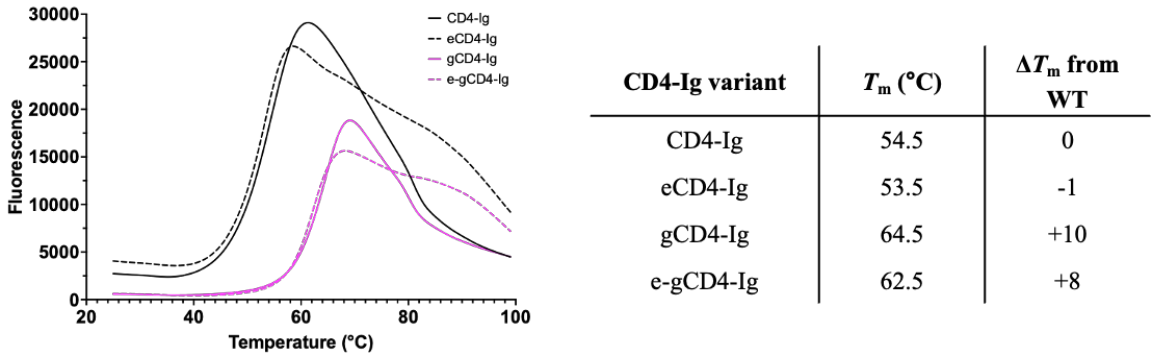


Extended Data Figure 1. Representative energy calculation of the D2 domain hydrophobics and DSF data. Shown is an example of an output from TRIAD using the PHOENIX scoring function and a sequence logo of the suggested amino acids at each position. The DSF data of nine variants is shown and their T_m s are reported in **Extended Data Table 1**.

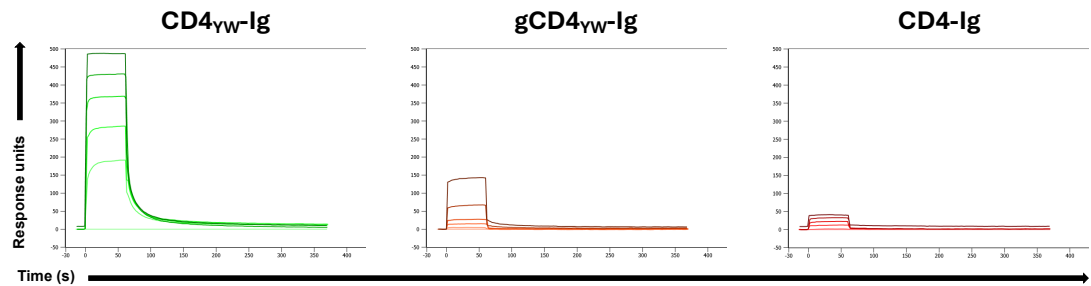


Extended Data Figure 2. Expression yield of CD4 D1D2 variants in Expi293 cells. Proteins were expressed in 20 mL volumes and purified as described in the Methods in technical triplicates. Yields were then calculated for one-liter expressions assuming a linear

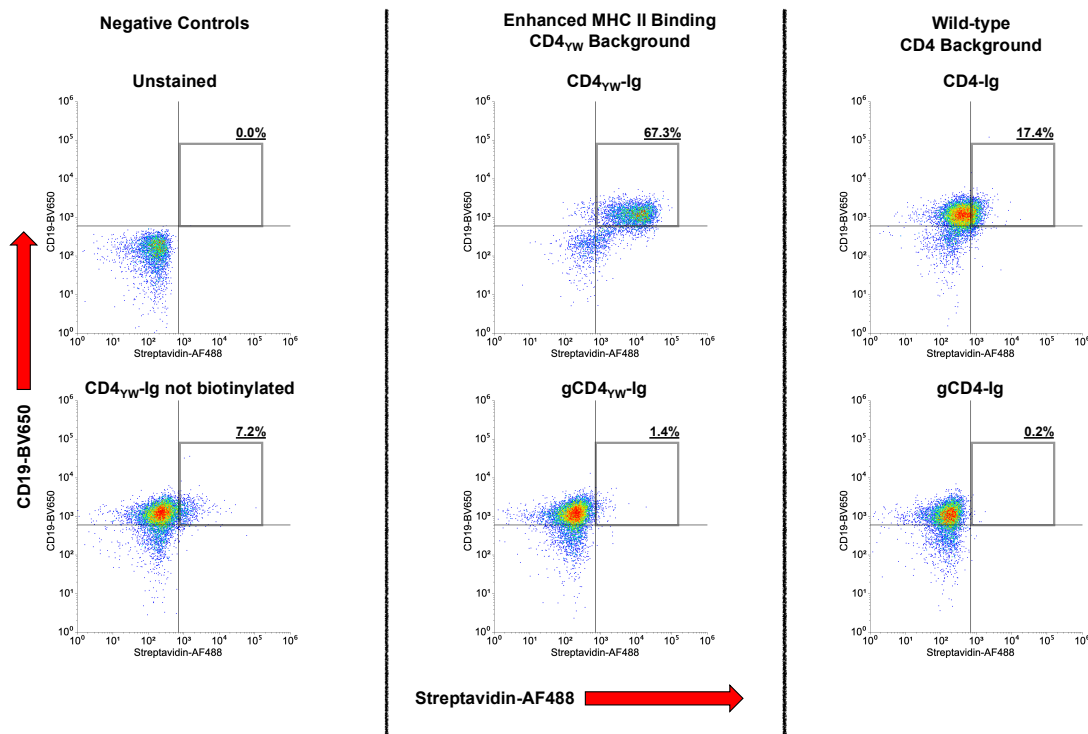
relationship with culture volume. gCD4 contains the VDKKKE substitutions and was expressed in the p3BNC plasmid. All other proteins were expressed in pTT5. Error bars are s.d.



Extended Data Figure 3. Apparent melting temperatures of CD4-Ig variants. Curves represent the average of technical triplicates.



Extended Data Figure 4. SPR of CD4-Ig variants against immobilized HLA-DR1. Proteins were injected for multi-cycle kinetics using 3-fold dilutions series from 50 μ M to 0.62 μ M as in Figure 2.



Extended Data Figure 5. Flow cytometry of CD4-Ig variants using tonsillar cells from donor #667. Flow cytometry was performed as in Figure 3.

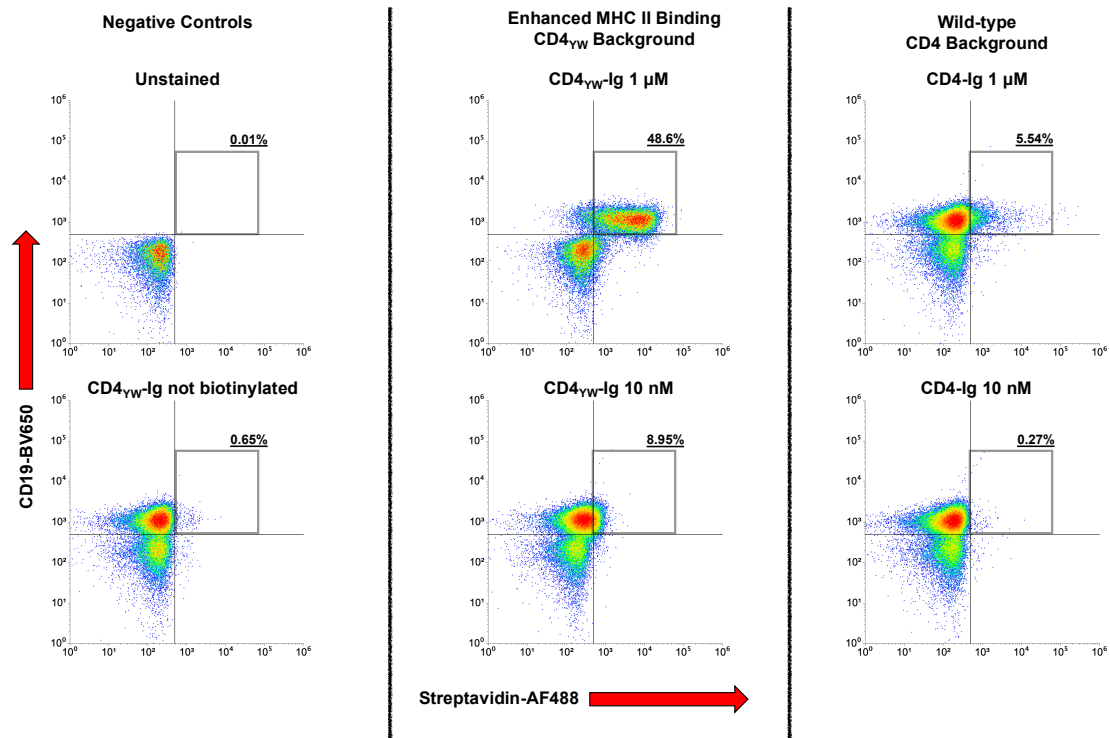
The amount of CD4-Ig delivered in a 15 mg/kg dose to a 70 kg adult male = $70\text{kg} * 15 \frac{\text{mg}}{\text{kg}} = 1.05\text{ g}$ of CD4-Ig

Assuming ~1.2 kg of leukocytes in a 70 kg adult male, the mass ratio of leukocytes to CD4-Ig = $\frac{1.2\text{kg}}{1.05\text{g}} \sim 1000 : 1$

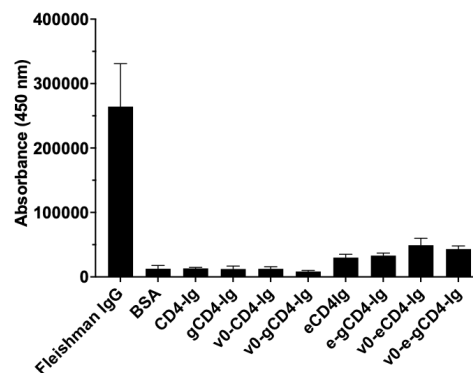
The mass of cells in 1 mL tonsil organoid cultures = 5,000,000 tonsil cells per well * $0.2 \frac{\text{ng}}{\text{lymphocyte}} = 1\text{ mg}$ of tonsil cells

The amount of CD4-Ig needed to replicate 1000 : 1 mass ratio = $\frac{1\text{ mg of tonsil cells}}{1000} = 1\text{ }\mu\text{g}$ of CD4-Ig

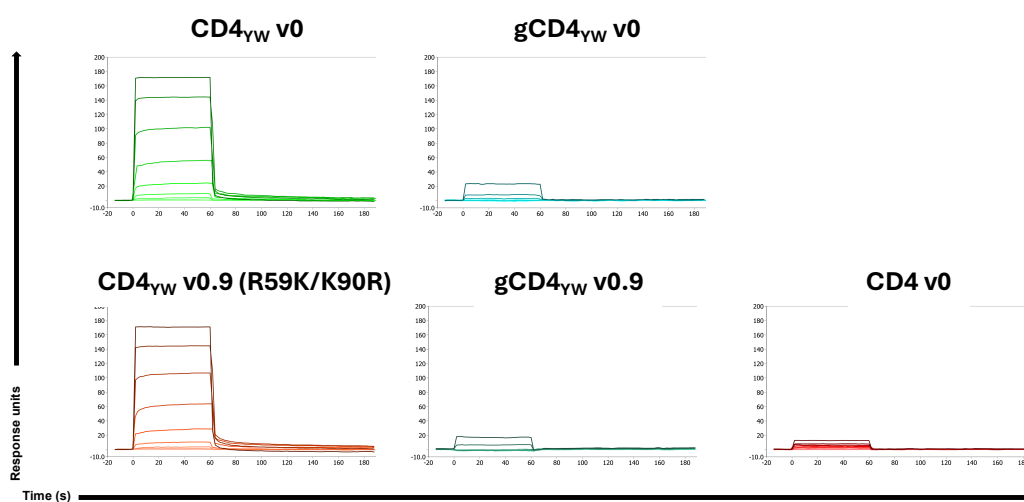
Extended Data Fig. 6: Calculation of the CD4-Ig amount needed to mimic clinical trial dosing in tonsil organoid cultures. Leukocyte mass was assumed using a reference 70 kg adult male (52).



Extended Data Figure 7. Flow cytometry of CD4-Ig variants at differing concentrations. In order to mimic clinical trial dosing in tonsil organoid cultures, CD4-Ig needed to be used at 10 nM (as determined in **Extended Data Fig. 6**). CD4_{YW}-Ig and CD4-Ig at 1 μ M and 10 nM were tested to determine which construct should be used in tonsil organoid cultures at 10 nM. Tonsil cells are from donor #667.



Extended Data Figure 8. Baculovirus polyreactivity assay of CD4-Ig variants. Proteins were assayed in quadruplicate. Fleischmann IgG was used as a positive control and BSA as a negative control. Error bars are s.d.



Extended Data Figure 9. SPR of new CD4 variants against HLA-DR1. SPR was performed using multi-cycle kinetics and a 3-fold dilution series from 50 μM to 0.023 μM as in **Figure 2**.

3.10 References

1. P. D. Kwong, R. Wyatt, J. Robinson, R. W. Sweet, J. Sodroski, W. A. Hendrickson, Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* **393**, 648–659 (1998).
2. X. X. Wang, Y. Li, Y. Yin, M. Mo, Q. Wang, W. Gao, L. Wang, R. A. Mariuzza, Affinity maturation of human CD4 by yeast surface display and crystal structure of a CD4–HLA-DR1 complex. *Proc. Natl. Acad. Sci.* **108**, 15960–15965 (2011).
3. D. J. Capon, S. M. Chamow, J. Mordenti, S. A. Marsters, T. Gregory, H. Mitsuya, R. A. Byrn, C. Lucas, F. M. Wurm, J. E. Groopman, S. Broder, D. H. Smith, Designing CD4 immunoadhesins for AIDS therapy. *Nature* **337**, 525–531 (1989).
4. A. Traunecker, J. Schneider, H. Kiefer, K. Karjalainen, Highly efficient neutralization of HIV with recombinant CD4-immunoglobulin molecules. *Nature* **339**, 68–70 (1989).
5. W. Chen, Y. Feng, P. Prabakaran, T. Ying, Y. Wang, J. Sun, C. D. S. Macedo, Z. Zhu, Y. He, V. R. Polonis, D. S. Dimitrov, Exceptionally Potent and Broadly Cross-Reactive, Bispecific Multivalent HIV-1 Inhibitors Based on Single Human CD4 and Antibody Domains. *J. Virol.* **88**, 1125–1139 (2014).
6. G. P. Allaway, K. L. Davis-Bruno, G. A. Beaudry, E. B. Garcia, E. L. Wong, A. M. Ryder, K. W. Hasel, M.-C. Gauduin, R. A. Koup, J. S. McDOUGAL, P. J. Maddon, Expression and Characterization of CD4-IgG2, a Novel Heterotetramer That Neutralizes Primary HIV Type 1 Isolates. *AIDS Res. Hum. Retroviruses* **11**, 533–539 (1995).
7. M. R. Gardner, L. M. Kattenhorn, H. R. Kondur, M. von Schaewen, T. Dorfman, J. J. Chiang, K. G. Haworth, J. M. Decker, M. D. Alpert, C. C. Bailey, E. S. Neale, C. H. Fellinger, V. R. Joshi, S. P. Fuchs, J. M. Martinez-Navio, B. D. Quinlan, A. Y. Yao, H. Mouquet, J. Gorman, B. Zhang, P. Poignard, M. C. Nussenzweig, D. R. Burton, P. D. Kwong, M. Piatak, J. D. Lifson, G. Gao, R. C. Desrosiers, D. T. Evans, B. H. Hahn, A. Ploss, P. M. Cannon, M. S. Seaman, M. Farzan, AAV-expressed eCD4-Ig provides durable protection from multiple SHIV challenges. *Nature* **519**, 87–91 (2015).
8. A. P. West, R. P. Galimidi, C. P. Foglesong, P. N. P. Gnanapragasam, J. S. Klein, P. J. Bjorkman, Evaluation of CD4-CD4i Antibody Architectures Yields Potent, Broadly Cross-Reactive Anti-Human Immunodeficiency Virus Reagents. *J. Virol.* **84**, 261–269 (2010).

9. I. Fetzter, M. R. Gardner, M. E. Davis-Gardner, N. R. Prasad, B. Alfant, J. A. Weber, M. Farzan, eCD4-Ig Variants That More Potently Neutralize HIV-1. *J. Virol.* **92**, e02011-17 (2018).
10. C. H. Fellingner, M. R. Gardner, J. A. Weber, B. Alfant, A. S. Zhou, M. Farzan, eCD4-Ig Limits HIV-1 Escape More Effectively than CD4-Ig or a Broadly Neutralizing Antibody. *J. Virol.* **93**, 10.1128/jvi.00443-19 (2019).
11. W. T. Shearer, R. J. Israel, S. Starr, C. V. Fletcher, D. Wara, M. Rathore, J. Church, J. DeVille, T. Fenton, B. Graham, P. Samson, S. Staprans, J. McNamara, J. Moye, P. J. Maddon, W. C. Olson, for the Pediatric AIDS Clinical Trials Group Protocol 351 Study Team, Recombinant CD4-IgG2 in Human Immunodeficiency Virus Type 1—Infected Children: Phase 1/2 Study. *J. Infect. Dis.* **182**, 1774–1779 (2000).
12. J. M. Jacobson, I. Lowy, C. V. Fletcher, T. J. O'Neill, D. N. H. Tran, T. J. Ketas, A. Trkola, M. E. Klotman, P. J. Maddon, W. C. Olson, R. J. Israel, Single-Dose Safety, Pharmacology, and Antiviral Activity of the Human Immunodeficiency Virus (HIV) Type 1 Entry Inhibitor PRO 542 in HIV-Infected Adults. *J. Infect. Dis.* **182**, 326–329 (2000).
13. A. Goyal, M. Gardner, B. T. Mayer, K. R. Jerome, M. Farzan, J. T. Schiffer, E. F. Cardozo-Ojeda, Estimation of the in vivo neutralization potency of eCD4Ig and conditions for AAV-mediated production for SHIV long-term remission. *Sci. Adv.* **8**, eabj5666 (2022).
14. A. Pan, C. C. Bailey, T. Ou, J. Xu, T. Aristotelous, X. Liu, B. Hu, G. Crynen, N. Skamangas, N. Bronkema, M. H. Tran, H. Mou, X. Zhang, M. D. Alpert, Y. Yin, M. Farzan, W. He, In vivo affinity maturation of the CD4 domains of an HIV-1-entry inhibitor. *Nat. Biomed. Eng.* **8**, 1715–1729 (2024).
15. C. V. Fletcher, J. G. DeVille, P. M. Samson, J. H. Moye, J. A. Church, H. M. L. Spiegel, P. Palumbo, T. Fenton, M. E. Smith, B. Graham, J. M. Krammer, W. T. Shearer, Nonlinear pharmacokinetics of high-dose recombinant fusion protein CD4-IgG2 (PRO 542) observed in HIV-1-infected children. *J. Allergy Clin. Immunol.* **119**, 747–750 (2007).
16. L. Corey, P. B. Gilbert, M. Juraska, D. C. Montefiori, L. Morris, S. T. Karuna, S. Edupuganti, N. M. Mgodi, A. C. deCamp, E. Rudnicki, Y. Huang, P. Gonzales, R. Cabello, C. Orrell, J. R. Lama, F. Laher, E. M. Lazarus, J. Sanchez, I. Frank, J. Hinojosa, M. E. Sobieszczyk, K. E. Marshall, P. G. Mukwekwerere, J. Makhema, L. R. Baden, J. I. Mullins, C. Williamson, J. Hural, M. J. McElrath, C. Bentley, S. Takuva,

- M. M. G. Lorenzo, D. N. Burns, N. Espy, A. K. Randhawa, N. Kochar, E. Piwowar-Manning, D. J. Donnell, N. Sista, P. Andrew, J. G. Kublin, G. Gray, J. E. Ledgerwood, J. R. Mascola, M. S. Cohen, Two Randomized Trials of Neutralizing Antibodies to Prevent HIV-1 Acquisition. *N. Engl. J. Med.* **384**, 1003–1014 (2021).
17. N. N. Mkhize, A. E. J. Yssel, H. Kaldine, R. T. van Dorsten, A. S. W. Davis, N. Beaume, D. Matten, B. Lambson, T. Modise, P. Kgagudi, T. York, D. H. Westfall, E. E. Giorgi, B. Korber, C. Anthony, R. E. Mapengo, V. Bekker, E. Domin, A. Eaton, W. Deng, A. DeCamp, Y. Huang, P. B. Gilbert, A. Gwashu-Nyangiwe, R. Thebus, N. Ndabambi, D. Mielke, N. Mgodi, S. Karuna, S. Edupuganti, M. S. Seaman, L. Corey, M. S. Cohen, J. Hural, M. J. McElrath, J. I. Mullins, D. Montefiori, P. L. Moore, C. Williamson, L. Morris, Neutralization profiles of HIV-1 viruses from the VRC01 Antibody Mediated Prevention (AMP) trials. *PLOS Pathog.* **19**, e1011469 (2023).
 18. J. F. Scheid, J. A. Horwitz, Y. Bar-On, E. F. Kreider, C.-L. Lu, J. C. C. Lorenzi, A. Feldmann, M. Braunschweig, L. Nogueira, T. Oliveira, I. Shimeliovich, R. Patel, L. Burke, Y. Z. Cohen, S. Hadrigan, A. Settler, M. Witmer-Pack, J. West, B. Juelg, T. Keler, T. Hawthorne, B. Zingman, R. M. Gulick, N. Pfeifer, G. H. Learn, M. S. Seaman, P. J. Bjorkman, F. Klein, S. J. Schlesinger, B. D. Walker, B. H. Hahn, M. C. Nussenzweig, M. Caskey, HIV-1 antibody 3BNC117 suppresses viral rebound in humans during treatment interruption. *Nature* **535**, 556–560 (2016).
 19. M. Caskey, T. Schoofs, H. Gruell, A. Settler, T. Karagounis, E. F. Kreider, B. Murrell, N. Pfeifer, L. Nogueira, T. Y. Oliveira, G. H. Learn, Y. Z. Cohen, C. Lehmann, D. Gillor, I. Shimeliovich, C. Unson-O'Brien, D. Weiland, A. Robles, T. Kümmerle, C. Wyen, R. Levin, M. Witmer-Pack, K. Eren, C. Ignacio, S. Kiss, A. P. West, H. Mouquet, B. S. Zingman, R. M. Gulick, T. Keler, P. J. Bjorkman, M. S. Seaman, B. H. Hahn, G. Fätkenheuer, S. J. Schlesinger, M. C. Nussenzweig, F. Klein, Antibody 10-1074 suppresses viremia in HIV-1-infected individuals. *Nat. Med.* **23**, 185–191 (2017).
 20. M. A. G. Hoffmann, Y. Bar-On, Z. Yang, H. B. Gristick, P. N. P. Gnanapragasam, J. Vielmetter, M. C. Nussenzweig, P. J. Bjorkman, Nanoparticles presenting clusters of CD4 expose a universal vulnerability of HIV-1 by mimicking target cells. *Proc. Natl. Acad. Sci.* **117**, 18719–18728 (2020).
 21. A. G. Dalgleish, P. C. L. Beverley, P. R. Clapham, D. H. Crawford, M. F. Greaves, R. A. Weiss, The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. *Nature* **312**, 763–767 (1984).

22. D. Klatzmann, E. Champagne, S. Chamaret, J. Gruest, D. Guetard, T. Hercend, J.-C. Gluckman, L. Montagnier, T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV. *Nature* **312**, 767–768 (1984).
23. H. Choe, M. Farzan, Y. Sun, N. Sullivan, B. Rollins, P. D. Ponath, L. Wu, C. R. Mackay, G. LaRosa, W. Newman, N. Gerard, C. Gerard, J. Sodroski, The β -Chemokine Receptors CCR3 and CCR5 Facilitate Infection by Primary HIV-1 Isolates. *Cell* **85**, 1135–1148 (1996).
24. T. Dragic, V. Litwin, G. P. Allaway, S. R. Martin, Y. Huang, K. A. Nagashima, C. Cayan, P. J. Maddon, R. A. Koup, J. P. Moore, W. A. Paxton, HIV-1 entry into CD4⁺ cells is mediated by the chemokine receptor CC-CKR-5. *Nature* **381**, 667–673 (1996).
25. H. Deng, R. Liu, W. Ellmeier, S. Choe, D. Unutmaz, M. Burkhart, P. D. Marzio, S. Marmon, R. E. Sutton, C. M. Hill, C. B. Davis, S. C. Peiper, T. J. Schall, D. R. Littman, N. R. Landau, Identification of a major co-receptor for primary isolates of HIV-1. *Nature* **381**, 661–666 (1996).
26. K.-M. A. Dam, C. Fan, Z. Yang, P. J. Bjorkman, Intermediate conformations of CD4-bound HIV-1 Env heterotrimers. *Nature* **623**, 1017–1025 (2023).
27. W. Li, Z. Qin, E. Nand, M. W. Grunst, J. R. Grover, J. W. Bess, J. D. Lifson, M. B. Zwick, H. D. Tagare, P. D. Uchil, W. Mothes, HIV-1 Env trimers asymmetrically engage CD4 receptors in membranes. *Nature* **623**, 1026–1033 (2023).
28. J. T. Ryman, B. Meibohm, Pharmacokinetics of Monoclonal Antibodies. *CPT Pharmacomet. Syst. Pharmacol.* **6**, 576–588 (2017).
29. A. D. McConnell, X. Zhang, J. L. Macomber, B. Chau, J. C. Sheffer, S. Rahmanian, E. Hare, V. Spasojevic, R. A. Horlick, D. J. King, P. M. Bowers, A general approach to antibody thermostabilization. *mAbs* **6**, 1274–1282 (2014).
30. G. An, Concept of Pharmacologic Target-Mediated Drug Disposition in Large-Molecule and Small-Molecule Compounds. *J. Clin. Pharmacol.* **60**, 149–163 (2020).
31. C. Doyle, J. L. Strominger, Interaction between CD4 and class II MHC molecules mediates cell adhesion. *Nature* **330**, 256–259 (1987).
32. B. Benacerraf, Role of MHC Gene Products in Immune Regulation. *Science* **212**, 1229–1238 (1981).

33. Y. Okura, Y. Ikawa-Teranishi, A. Mizoroki, N. Takahashi, T. Tsushima, M. Irie, Z. Harfuddin, M. Miura-Okuda, S. Ito, G. Nakamura, H. Takesue, Y. Ozono, M. Nishihara, K. Yamada, S. W. Gan, A. Hayasaka, S. Ishii, T. Wakabayashi, M. Muraoka, N. Nagaya, H. Hino, T. Nemoto, T. Kuramochi, T. Torizawa, H. Shimada, T. Kitazawa, M. Okazaki, J. Nezu, L. M. Sollid, T. Igawa, Characterizations of a neutralizing antibody broadly reactive to multiple gluten peptide:HLA-DQ2.5 complexes in the context of celiac disease. *Nat. Commun.* **14**, 8502 (2023).
34. R. M. Russell, F. Bibollet-Ruche, W. Liu, S. Sherrill-Mix, Y. Li, J. Connell, D. E. Loy, S. Trimboli, A. G. Smith, A. N. Avitto, M. V. P. Gondim, L. J. Plenderleith, K. S. Wetzel, R. G. Collman, A. Ayoub, A. Esteban, M. Peeters, W. J. Kohler, R. A. Miller, S. François-Souquiere, W. M. Switzer, V. M. Hirsch, P. A. Marx, A. K. Piel, F. A. Stewart, A. V. Georgiev, V. Sommer, P. Bertolani, J. A. Hart, T. B. Hart, G. M. Shaw, P. M. Sharp, B. H. Hahn, CD4 receptor diversity represents an ancient protection mechanism against primate lentiviruses. *Proc. Natl. Acad. Sci.* **118**, e2025914118 (2021).
35. D. E. Knuth, *The Art of Computer Programming* (Addison-Wesley, Reading, Mass, 3rd ed., 1997).
36. Schrodinger, L. (2010) The PyMOL Molecular Graphics System, Version 1.3r1. - References - Scientific Research Publishing. <https://www.scirp.org/reference/ReferencesPapers?ReferenceID=1571978>.
37. A. Ashkenazi, L. G. Presta, S. A. Marsters, T. R. Camerato, K. A. Rosenthal, B. M. Fendly, D. J. Capon, Mapping the CD4 binding site for human immunodeficiency virus by alanine-scanning mutagenesis. *Proc. Natl. Acad. Sci.* **87**, 7150–7154 (1990).
38. E. G. Baker, G. J. Bartlett, M. P. Crump, R. B. Sessions, N. Linden, C. F. J. Faul, D. N. Woolfson, Local and macroscopic electrostatic interactions in single α -helices. *Nat. Chem. Biol.* **11**, 221–228 (2015).
39. M. N. Rushdi, V. Pan, K. Li, H.-K. Choi, S. Travaglino, J. Hong, F. Griffiths, P. Agnihotri, R. A. Mariuzza, Y. Ke, C. Zhu, Cooperative binding of T cell receptor and CD4 to peptide-MHC enhances antigen sensitivity. *Nat. Commun.* **13**, 7055 (2022).
40. P. Jönsson, J. H. Southcombe, A. M. Santos, J. Huo, R. A. Fernandes, J. McColl, M. Lever, E. J. Evans, A. Hudson, V. T. Chang, T. Hanke, A. Godkin, P. D. Dunne, M. H. Horrocks, M. Palayret, G. R. Screaton, J. Petersen, J. Rossjohn, L. Fugger, O. Dushek, X.-N. Xu, S. J. Davis, D. Klennerman, Remarkably low affinity of CD4/peptide-major

- histocompatibility complex class II protein interactions. *Proc. Natl. Acad. Sci.* **113**, 5682–5687 (2016).
41. H. H. Nguyen, J. Park, S. Kang, M. Kim, Surface Plasmon Resonance: A Versatile Technique for Biosensor Applications. *Sensors* **15**, 10481–10510 (2015).
 42. H. Wu, P. D. Kwong, W. A. Hendrickson, Dimeric association and segmental variability in the structure of human CD4. *Nature* **387**, 527–530 (1997).
 43. R. F. Alford, A. Leaver-Fay, J. R. Jeliazkov, M. J. O'Meara, F. P. DiMaio, H. Park, M. V. Shapovalov, P. D. Renfrew, V. K. Mulligan, K. Kappel, J. W. Labonte, M. S. Pacella, R. Bonneau, P. Bradley, R. L. Jr. Dunbrack, R. Das, D. Baker, B. Kuhlman, T. Kortemme, J. J. Gray, The Rosetta All-Atom Energy Function for Macromolecular Modeling and Design. *J. Chem. Theory Comput.* **13**, 3031–3048 (2017).
 44. Y. T. Tang, G. R. Marshall, PHOENIX: A Scoring Function for Affinity Prediction Derived Using High-Resolution Crystal Structures and Calorimetry Measurements. *J. Chem. Inf. Model.* **51**, 214–228 (2011).
 45. F. H. Niesen, H. Berglund, M. Vedadi, The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* **2**, 2212–2221 (2007).
 46. S. M. Schnittman, H. C. Lane, J. Roth, A. Burrows, T. M. Folks, J. H. Kehrl, S. Koenig, P. Berman, A. S. Fauci, Characterization of GP120 binding to CD4 and an assay that measures ability of sera to inhibit this binding. *J. Immunol.* **141**, 4181–4186 (1988).
 47. D. G. Myszka, R. W. Sweet, P. Hensley, M. Brigham-Burke, P. D. Kwong, W. A. Hendrickson, R. Wyatt, J. Sodroski, M. L. Doyle, Energetics of the HIV gp120-CD4 binding reaction. *Proc. Natl. Acad. Sci.* **97**, 9026–9031 (2000).
 48. G. Wozniak-Knopp, J. Stadlmann, F. Rüker, Stabilisation of the Fc Fragment of Human IgG1 by Engineered Intradomain Disulfide Bonds. *PLOS ONE* **7**, e30083 (2012).
 49. L. E. Wagar, A. Salahudeen, C. M. Constantz, B. S. Wendel, M. M. Lyons, V. Mallajosyula, L. P. Jatt, J. Z. Adamska, L. K. Blum, N. Gupta, K. J. L. Jackson, F. Yang, K. Röltgen, K. M. Roskin, K. M. Blaine, K. D. Meister, I. N. Ahmad, M. Cortese, E. G. Dora, S. N. Tucker, A. I. Sperling, A. Jain, D. H. Davies, P. L. Felgner, G. B. Hammer, P. S. Kim, W. H. Robinson, S. D. Boyd, C. J. Kuo, M. M. Davis, Modeling human adaptive immune responses with tonsil organoids. *Nat. Med.* **27**, 125–135 (2021).

50. A. Casasola-LaMacchia, M. S. Ritorto, R. J. Seward, N. Ahyi-Amendah, A. Ciarla, T. P. Hickling, H. Neubert, Human leukocyte antigen class II quantification by targeted mass spectrometry in dendritic-like cell lines and monocyte-derived dendritic cells. *Sci. Rep.* **11**, 1028 (2021).
51. D. Beckett, E. Kovaleva, P. J. Schatz, A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein Sci. Publ. Protein Soc.* **8**, 921–929 (1999).
52. R. Sender, Y. Weiss, Y. Navon, I. Milo, N. Azulay, L. Keren, S. Fuchs, D. Ben-Zvi, E. Noor, R. Milo, The total mass, number, and distribution of immune cells in the human body. *Proc. Natl. Acad. Sci.* **120**, e2308511120 (2023).
53. K. L. Otipoby, Y. Sasaki, M. Schmidt-Supprrian, A. Patke, R. Gareus, M. Pasparakis, A. Tarakhovsky, K. Rajewsky, BAFF activates Akt and Erk through BAFF-R in an IKK1-dependent manner in primary mouse B cells. *Proc. Natl. Acad. Sci.* **105**, 12435–12438 (2008).
54. A. deCamp, P. Hraber, R. T. Bailer, M. S. Seaman, C. Ochsenbauer, J. Kappes, R. Gottardo, P. Edlefsen, S. Self, H. Tang, K. Greene, H. Gao, X. Daniell, M. Sarzotti-Kelsoe, M. K. Gorny, S. Zolla-Pazner, C. C. LaBranche, J. R. Mascola, B. T. Korber, D. C. Montefiori, Global Panel of HIV-1 Env Reference Strains for Standardized Assessments of Vaccine-Elicited Neutralizing Antibodies. *J. Virol.* **88**, 2489–2507 (2014).
55. I. Hötzel, F.-P. Theil, L. J. Bernstein, S. Prabhu, R. Deng, L. Quintana, J. Lutman, R. Sibia, P. Chan, D. Bumbaca, P. Fielder, P. J. Carter, R. F. Kelley, A strategy for risk mitigation of antibodies with fast clearance. *mAbs* **4**, 753–760 (2012).
56. P. R. Hinton, J. M. Xiong, M. G. Johlfs, M. T. Tang, S. Keller, N. Tsurushita, An Engineered Human IgG1 Antibody with Longer Serum Half-Life. *J. Immunol.* **176**, 346–356 (2006).
57. S. H. Tam, S. G. McCarthy, K. Brosnan, K. M. Goldberg, B. J. Scallon, Correlations between pharmacokinetics of IgG antibodies in primates vs. FcRn-transgenic mice reveal a rodent model with predictive capabilities. *mAbs* **5**, 397–405 (2013).
58. G. C. Bosma, R. P. Custer, M. J. Bosma, A severe combined immunodeficiency mutation in the mouse. *Nature* **301**, 527–530 (1983).
59. M. Wu, S. Hussain, Y.-H. He, R. Pasula, P. A. Smith, W. J. Martin, Genetically engineered macrophages expressing IFN- γ restore alveolar immune function in scid mice. *Proc. Natl. Acad. Sci.* **98**, 14589–14594 (2001).

60. M. M. Sajadi, A. Abbasi, Z. R. Tehrani, C. Siska, R. Clark, W. Chi, M. S. Seaman, D. Mielke, K. Wagh, Q. Liu, T. Jumpa, R. R. Ketchum, D. N. Nguyen, W. D. Tolbert, B. G. Pierce, B. Atkinson, D. Deming, M. Sprague, A. Asakawa, D. Ferrer, Y. Dunn, S. Calvillo, R. Yin, J. D. Guest, B. Korber, B. T. Mayer, A. H. Sato, X. Ouyang, S. Foulke, P. Habibzadeh, M. Karimi, A. Aslanabadi, M. Hojabri, S. Saadat, R. Zareidoodeji, M. Kędzior, E. Pozharski, A. Heredia, H. Chen, D. Montefiori, G. Ferrari, M. Pazgier, G. K. Lewis, J. G. Jardine, P. Lusso, A. DeVico, A comprehensive engineering strategy improves potency and manufacturability of a near pan-neutralizing antibody against HIV. *Structure* **33**, 1150-1164.e8 (2025).
61. Y. D. Kwon, A. Pegu, E. S. Yang, B. Zhang, M. F. Bender, M. Asokan, Q. Liu, K. McKee, B. C. Lin, T. Liu, M. K. Louder, R. Rawi, M. Reveiz, A. J. Schaub, C.-H. Shen, N. A. Doria-Rose, P. Lusso, J. R. Mascola, P. D. Kwong, Improved pharmacokinetics of HIV-neutralizing VRC01-class antibodies achieved by reduction of net positive charge on variable domain. *mAbs* **15**, 2223350 (2023).
62. R. E. Hussey, N. E. Richardson, M. Kowalski, N. R. Brown, H.-C. Chang, R. F. Siliciano, T. Dorfman, B. Walker, J. Sodroski, E. L. Reinherz, A soluble CD4 protein selectively inhibits HIV replication and syncytium formation. *Nature* **331**, 78–81 (1988).
63. M. E. Davis-Gardner, M. R. Gardner, B. Alfant, M. Farzan, eCD4-Ig promotes ADCC activity of sera from HIV-1-infected patients. *PLOS Pathog.* **13**, e1006786 (2017).
64. D. J. DiLillo, J. V. Ravetch, Differential Fc-Receptor Engagement Drives an Anti-tumor Vaccinal Effect. *Cell* **161**, 1035–1045 (2015).
65. A. Ardeshir, D. O'Hagan, I. Mehta, S. Shandilya, L. L. J. Hopkins, L. Adamson, M. J. Kuroda, P. A. Hahn, L. A. B. da Costa, S. P. Fuchs, J. M. Martinez-Navio, M. R. Gardner, K. K. A. Van Rompay, D. M. Magnani, J. D. Lifson, G. Gao, M. Farzan, R. C. Desrosiers, J. Das, M. A. Martins, Determinants of successful AAV-vectored delivery of HIV-1 bNAbs in early life. *Nature* **645**, 1020–1028 (2025).
66. D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison, H. O. Smith, Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
67. T. Zhou, L. Xu, B. Dey, A. J. Hessel, D. Van Ryk, S.-H. Xiang, X. Yang, M.-Y. Zhang, M. B. Zwick, J. Arthos, D. R. Burton, D. S. Dimitrov, J. Sodroski, R. Wyatt, G. J. Nabel, P. D. Kwong, Structural definition of a conserved neutralization epitope on HIV-1 gp120. *Nature* **445**, 732–737 (2007).

68. M. V. Shapovalov, R. L. Dunbrack, A Smoothed Backbone-Dependent Rotamer Library for Proteins Derived from Adaptive Kernel Density Estimates and Regressions. *Structure* **19**, 844–858 (2011).
69. C. Fan, A. A. Cohen, M. Park, A. F.-H. Hung, J. R. Keeffe, P. N. P. Gnanapragasam, Y. E. Lee, H. Gao, L. M. Kakutani, Z. Wu, H. Kleanthous, K. E. Malecek, J. C. Williams, P. J. Bjorkman, Neutralizing monoclonal antibodies elicited by mosaic RBD nanoparticles bind conserved sarbecovirus epitopes. *Immunity* **55**, 2419-2435.e10 (2022).
70. C. O. Barnes, T. Schoofs, P. N. P. Gnanapragasam, J. Golijanin, K. E. Huey-Tubman, H. Gruell, P. Schommers, N. Suh-Toma, Y. E. Lee, J. C. Cetrulo Lorenzi, A. Piechocka-Trocha, J. F. Scheid, A. P. West, B. D. Walker, M. S. Seaman, F. Klein, M. C. Nussenzweig, P. J. Bjorkman, A naturally arising broad and potent CD4-binding site antibody with low somatic mutation. *Sci. Adv.* **8**, eabp8155 (2022).
71. D. C. Montefiori, Evaluating Neutralizing Antibodies Against HIV, SIV, and SHIV in Luciferase Reporter Gene Assays. *Curr. Protoc. Immunol.* **64**, 12.11.1-12.11.17 (2004).
72. D. C. Montefiori, “Measuring HIV Neutralization in a Luciferase Reporter Gene Assay” in *HIV Protocols*, V. R. Prasad, G. V. Kalpana, Eds. (Humana Press, Totowa, NJ, 2009; https://doi.org/10.1007/978-1-59745-170-3_26), pp. 395–405.
73. H. B. Gristick, H. Hartweger, M. Loewe, J. van Schooten, V. Ramos, T. Y. Oliveira, Y. Nishimura, N. S. Koranda, A. Wall, K.-H. Yao, D. Poston, A. Gazumyan, M. Wiatr, M. Horning, J. R. Keeffe, M. A. G. Hoffmann, Z. Yang, M. E. Abernathy, K.-M. A. Dam, H. Gao, P. N. P. Gnanapragasam, L. M. Kakutani, A. J. Pavlovitch-Bedzyk, M. S. Seaman, M. Howarth, A. T. McGuire, L. Stamatatos, M. A. Martin, A. P. West, M. C. Nussenzweig, P. J. Bjorkman, CD4 binding site immunogens elicit heterologous anti-HIV-1 neutralizing antibodies in transgenic and wild-type animals. *Sci. Immunol.* **8**, eade6364 (2023).
74. A. P. West, L. Scharf, J. Horwitz, F. Klein, M. C. Nussenzweig, P. J. Bjorkman, Computational analysis of anti-HIV-1 antibody neutralization panel data to identify potential functional epitope residues. *Proc. Natl. Acad. Sci.* **110**, 10598–10603 (2013).

TOWARDS DISCOVERING NANOBODIES THAT TARGET INVARIANT DOMAINS IN HIV-1

4.1 Summary

A conserved epitope on the HIV-1 Env trimer that is exposed only after engagement with its host receptor target CD4, termed the caldera, represents a potentially vulnerable site for compact binders such as V_{HH} nanobodies. To identify nanobodies targeting this region, we carried out an initial screening campaign using soluble Env trimers, termed SOSIPs, from multiple HIV-1 clades in a yeast display workflow with iterative magnetic- and fluorescence-based enrichment. Conformationally fixed and flexible SOSIPs enabled depletion of off-target binders and enrichment for binders to CD4-induced conformations, respectively. Parallel selections incorporating 21c scFv were used to reduce recovery of binders to a known CD4-induced site and bias discovery toward alternative open-state epitopes, including the caldera. This campaign yielded five candidates for expression and ELISA-based characterization. Although these candidates showed selective binding to open BG505-sCD4 relative to closed, prefusion DS-BG505, they also bound sCD4 alone and showed little-to-no binding to gp120 monomer, indicating recognition of a composite quaternary epitope formed by open Env and CD4. This epitope appears to be distinct from both the caldera and previously characterized CD4-induced sites, representing a novel target for potential HIV-1 entry inhibition. Thus, while this initial campaign did not isolate caldera-specific nanobodies, it enriched for candidates consistent with the screen design, established a screening framework for future discovery efforts, and revealed key design constraints for screening challenging targets.

4.2 Introduction

Despite major advances in antiretroviral therapy (ART) and antibody-based prevention, HIV-1 remains a major global health challenge because its extraordinary genetic and antigenic diversity limits the breadth of many interventions. Broadly neutralizing antibodies (bNAbs) such as 3BNC117, 10-1074, and VRC01 can suppress viremia, delay viral rebound, and protect against sensitive viral strains in prevention settings (1–5). In particular, the Antibody Mediated Prevention (AMP) trials demonstrated that VRC01 conferred significant protection (~75% efficacy) against susceptible HIV-1 strains, but provided no protection against resistant strains, establishing that bNAb-mediated prevention is contingent on *in vitro* viral susceptibility (1, 6–8). These limitations motivate efforts to identify new antibody-based therapeutics directed against the highly conserved, functionally constrained surfaces on the HIV-1 envelope (Env) glycoprotein, the sole protein accessible on the surface of the virus.

HIV-1 entry is mediated by Env, a trimer of gp120-gp41 heterodimers that first engages the host receptor CD4 alongside a co-receptor, CCR5 or CXCR4, to trigger membrane fusion (9–11). In the closed, prefusion trimer, many of the most conserved regions of Env are shielded by protomer packing and a shield of glycans covering most of its surface, greatly limiting antibody accessibility (12, 13). Upon CD4 engagement, however, Env undergoes large conformational rearrangements in which gp120 protomers rotate outward, thereby exposing previously occluded surfaces (11, 14–18). These receptor-induced states are central to viral entry and may therefore reveal conserved vulnerabilities that are less permissive to mutational escape by the virus.

We refer to one such CD4-induced surface as the caldera, a conserved, glycan-free region exposed near the center of the open trimer between gp120 protomers atop gp41 (**Figure 1**). This epitope is an attractive target because it is functionally constrained by the membrane fusion machinery and is more conserved than many antibody-accessible epitopes on the surface of closed Env. At the same time, the caldera is likely a challenging target, as it is transiently exposed, structurally dynamic, and positioned in a recessed region between

protomers. These features raise the possibility that compact binders may be better suited than conventional antibodies to target this region.

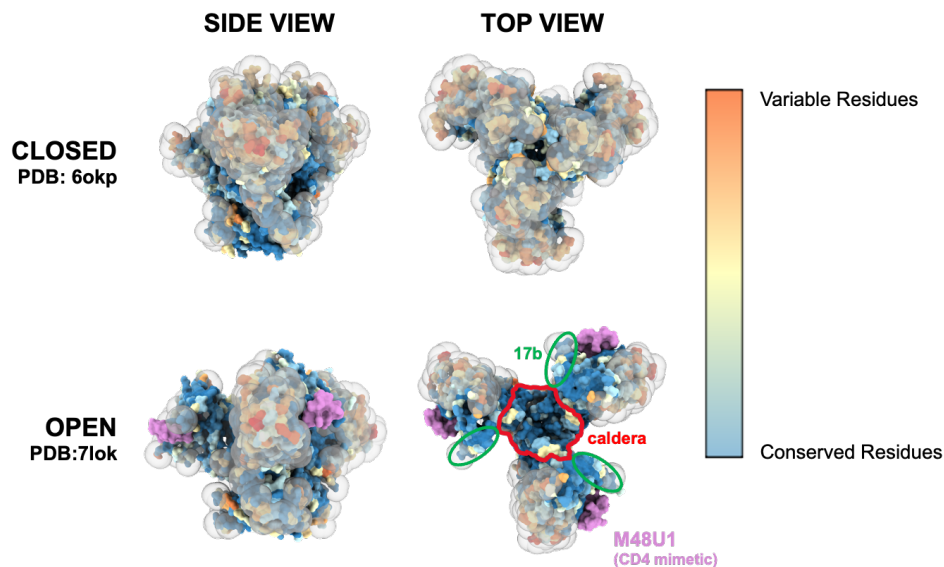


Figure 1. Heat map on HIV-1 Env showing residue conservation across isolated sequences.

Open and closed conformations shown, with the CD4-induced (CD4i) and caldera epitopes highlighted in green and red, respectively, in open Env. 17b IgG binding region used to show CD4i epitope.

Camelid single-domain antibodies (V_{HH} , or nanobodies) are well-suited for discovery efforts against such cryptic epitopes (19, 20). Their small size, single-domain architecture, and long CDR3 loops can enable access to recessed, sterically restricted epitopes that are difficult for conventional antibodies to reach. Nanobodies are also highly stable and can be readily reformatted into multivalent or multi-specific molecules, making them attractive for both mechanistic studies and therapeutic development. Their straightforward fusion to an IgG Fc also allows them to engage Fc-dependent effector functions, further expanding their use as antibody-based therapies. These properties suggest that nanobodies might provide a practical route toward targeting the HIV-1 caldera.

To discover caldera-directed binders, we developed a screening strategy based on soluble native-like Env trimers, termed SOSIPs, from multiple HIV-1 clades in both conformationally flexible and disulfide-stabilized forms (12, 21–24). Flexible open SOSIPs in the presence of soluble CD4 (sCD4) were used to enrich for binders recognizing receptor-induced open conformations, whereas disulfide-stabilized (DS) trimers were used to deplete binders recognizing closed, prefusion Env. We combined these engineered probes with a synthetic yeast display V_{HH} library and iterative rounds of magnetic- and fluorescence-based enrichment to favor the isolation of nanobodies recognizing CD4-induced conformations while reducing enrichment of off-target binders, including those targeting the CD4-induced (CD4i) epitope at the apex of gp120 protomers also revealed upon CD4 engagement. This first screening campaign yielded candidates with selective binding to open Env, though downstream characterization indicated recognition of a composite quaternary epitope involving both CD4 and open trimer rather than the caldera region alone. These results nonetheless establish a framework for future caldera-focused discovery efforts and identify design constraints that should improve subsequent screens.

4.3 Results

4.3.1 Production and validation of SOSIP Env trimers

To screen for nanobodies that bind selectively to CD4-induced epitopes, we prepared multiple soluble HIV-1 Env reagents representing three distinct HIV-1 clades for breadth. We selected BG505, B41, and ZM197M from clades A, B, and C, respectively, with SOSIPv2 stabilizing mutations (A501C-T605C, I559P, R6, ΔMPER) to retain native-like conformational flexibility (24). To generate conformationally-locked SOSIPs for negative selection, we introduced the 201C-433C disulfide (DS) bond into each to create DS-BG505, DS-B41, and DS-ZM197M (23). These DS-SOSIPs are unable to undergo CD4-induced conformational changes and therefore serve as tools to deplete nanobodies that bind to the exterior of the closed trimer. All six constructs were expressed in mammalian cells and purified over a PGT145 affinity column followed by size-exclusion chromatography (SEC) (Figure 2a). We confirmed by SDS-PAGE all six preps were clean and at the appropriate molecular weights (Figure 2b).

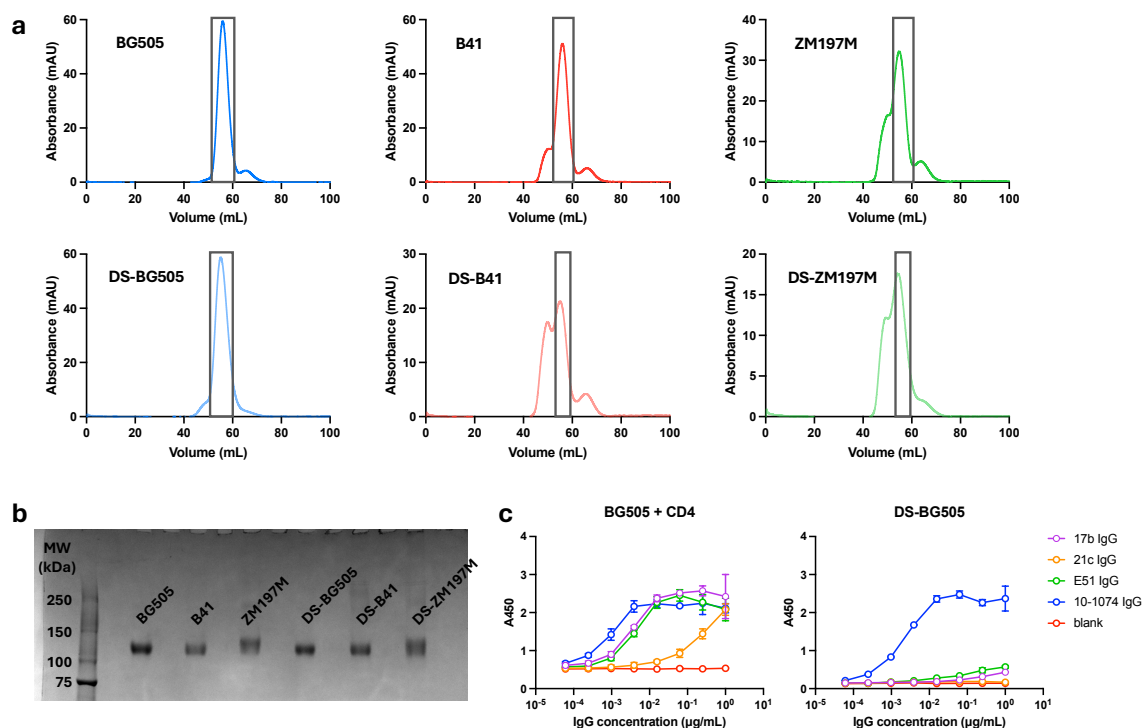


Figure 2. HIV Env SOSIP purification and characterization.

- (a) SEC traces after purification by Superose 6 Increase 10/300 GL.
- (b) Non-reduced SDS-PAGE analysis of pooled fractions shown for all six constructs.
- (c) ELISA binding data indicating minimal binding of CD4i IgGs (17b, 21c, E51) in closed DS-BG505 compared to open BG505 with CD4 (25–27). 10-1074 IgG is a positive control in both.

For downstream screening and binding assays, SOSIPs were modified using EZ-Link NHS-esters to randomly label primary amines on accessible lysine residues and the N-terminus. Using these reagents, we generated biotinylated and Alexa Fluor 647 (AF647)-labeled SOSIPs. Non-reduced SDS-PAGE on SOSIP-streptavidin complexes confirmed successful biotinylation for all six constructs via a shift in molecular weight. Enzyme-linked immunosorbent assay (ELISA) on streptavidin-coated plates using a panel of bNAbs confirmed epitope accessibility, including the CD4-induced (CD4i) epitope in non-DS constructs in the presence of soluble CD4 (sCD4). In the absence of sCD4, limited CD4i binding was observed in B41 and ZM197M SOSIPs, consistent with what has been

previously reported (24). DS-SOSIPs exhibited minimal binding at the CD4i epitope, as expected, highlighting the importance of using these conformationally-locked constructs to avoid inadvertently depleting candidates bound to partially open SOSIPs (**Figure 2c**). Fluorescence-activated cell sorting (FACS) using yeast displaying either the CD4bs nanobody J3 or the CD4i single-chain variable fragment (scFv) 17b likewise confirmed successful AF647 labeling and epitope accessibility for all six AF647-labeled constructs.

4.3.2 Screening for caldera-binding nanobodies using yeast display

Using a V_{HH} nanobody library developed for screening in yeast by the Kruse lab at Harvard, we performed iterative rounds of magnetic-activated cell sorting (MACS) using anti-AF647 magnetic beads against our AF647-labeled constructs to enrich for likely binders before performing FACS (28). MACS enrichment consisted of three rounds, M1-M3, with yeast cultures expanded in between each (**Figure 3, top row**).

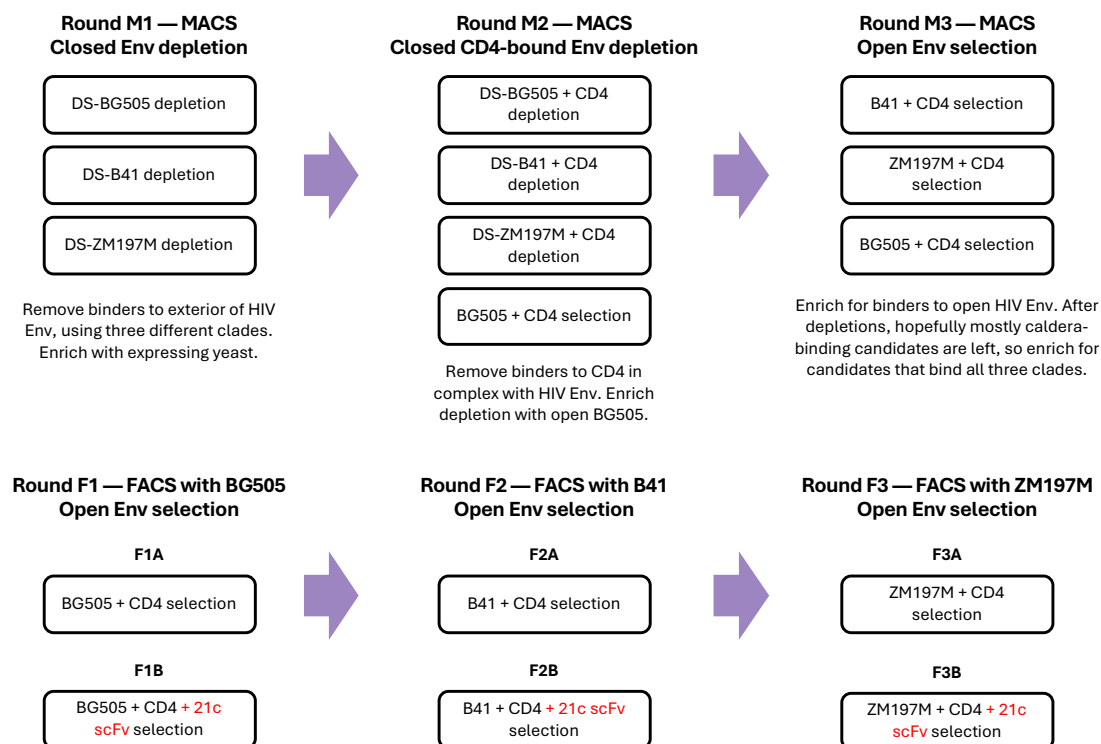


Figure 3. Screening workflow, including MACS (top row) and FACS (bottom row).

Iterative rounds of MACS (top row) and FACS (bottom row) to enrich for binders exclusive to open Env.

In the first MACS round, nanobodies were depleted using DS-SOSIPs to remove binders to closed Env. After incubating with DS-BG505, nanobody-displaying yeast were washed to remove excess unbound antigen before incubating with anti-AF647 magnetic beads. Another wash was performed to remove excess beads before running the culture through an LD depletion column by gravity at 4°C, with DS-BG505 binders being retained on the column. The flowthrough was then incubated with DS-B41 and the process repeated, including a third time with DS-ZM197M. For round two, DS-SOSIPs were complexed with a molar excess amount of sCD4 (1:6 DS-SOSIP:sCD4) before incubating with nanobody-displaying yeast to remove potential off-target binders to sCD4, including interfaces overlapping both sCD4 and DS-SOSIP. Despite not undergoing conformational changes, DS-SOSIPs are known to complex with sCD4 (29). Like in round one, excess antigen and excess beads were removed before collecting the flowthrough from LD columns. This was done with DS-BG505-sCD4, DS-B41-sCD4, and DS-ZM197M-sCD4 in succession before running a final selection using pre-incubated BG505-sCD4 (i.e., open Env in 1:6 BG505:sCD4 molar amounts) through an LS selection column to enrich for possible binders. In this case, our desired cells were eluted from the column before expanding. Finally, we ran a series of selection steps to enrich for open Env binders using our BG505, B41, and ZM197M antigens in succession, each pre-complexed with sCD4 (1:6 SOSIP:sCD4), through LS columns. The final enriched culture was expanded once more before FACS experiments.

FACS enrichment was similarly done over three rounds, F1-F3, with each based on binding to a single SOSIP antigen complexed with sCD4 (**Figure 3, bottom row**). In addition to our AF647-labeled SOSIP probes, each round included AF488-labeled anti-HA antibodies that bound to HA tags fused at the C-terminal ends of displayed nanobodies to sort between yeast that failed to display (no signal), yeast that displayed non-binders (488-positive), and yeast that displayed binders (488- and 647-double-positive). Additionally, each FACS round was split into two groups, A and B, with group B including 21c scFv in antigen pre-incubation to screen with SOSIP-sCD4-21c complexes. 21c is a CD4i antibody that binds exclusively to open Env with an epitope footprint that includes CD4, making the antibody lean “outward” from the center (26). In including it, we intended to block this epitope to deplete off-target

binders against this site, leaving only caldera-binding candidates in our double-positive cell population.

After the first FACS round with BG505-sCD4 (F1A) and BG505-sCD4-21c (F1B), we noticed the latter had a smaller double-positive population (1.86%, vs 5.86%), suggesting our second group indeed depleted additional (presumably off-target) binders (**Figure 4, left**). After the second round using our B41 SOSIP probe, the group with 21c (F2B) had an approximate 5-fold increase in the proportion of cells in the double-positive population to 10.14% (**Figure 4, center**). This jump was due to our enrichment in the previous round—while the double-positive population in the first group remained at ~5% (F1A and F2A), the double-positive population in the group with 21c saw an enrichment that excluded CD4i binders (F1B and F2B). With non-CD4i binders taking up a larger proportion of counts in the second group with 21c, we expect to see this jump in F2B but not F1B. Finally, the third round was done with ZM197M (**Figure 4, right**). For both groups, F3A and F3B, we had incredibly low double-positive counts (<1%), which we attribute to the relatively poor stability of our ZM197M probe, especially since the double-positive population was the only one with a drastic change. Cell stocks collected after MACS and each FACS run were deep sequenced to yield ~1000 unique sequences across all enrichment steps. Candidates with the highest sequence counts in each group were conservatively chosen across all cell stocks for a total of 31 candidates, with many showing clear patterns of enrichment (i.e., consistently high sequence counts after each enrichment step).

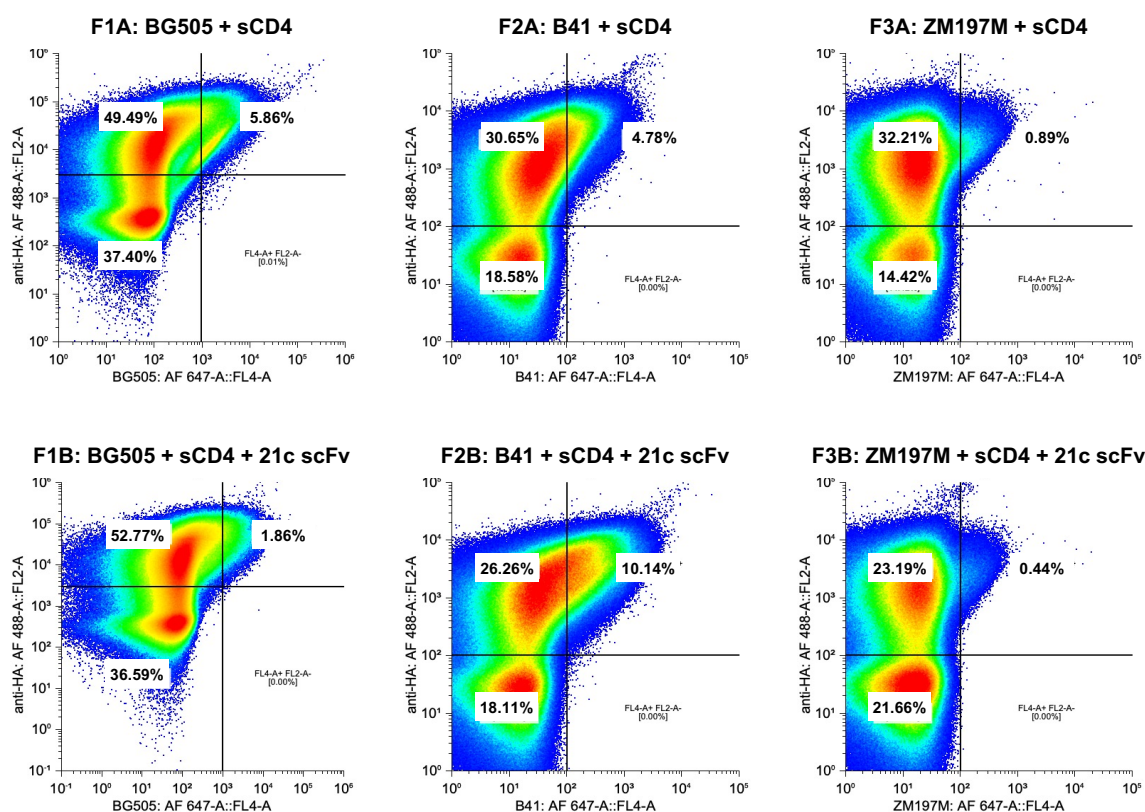


Figure 4. Iterative rounds of FACS enrichment using Env SOSIPs across three clades. BG505 (left), B41 (center), and ZM197M (right). Bottom row screens included 21c scFv.

Parallel efforts with industry collaborators at Protabit yielded additional candidates. Using a promiscuous pre-existing V_{HH} library they designed to screen for binders against vascular endothelial growth factor (VEGF), they executed a simplified protocol using only BG505 and DS-BG505, the most stable probes. MACS was first performed to deplete binders to closed DS-BG505 (with and without sCD4 to deplete DS-BG505-sCD4 binders) and followed up with a selection to enrich for binders against open BG505-sCD4. The enriched population was then further sorted through iterative rounds of FACS. Similar to MACS, the first FACS experiments involved depletion and selection rounds using DS-BG505 and BG505-sCD4, respectively. A second round was performed with similar depletion and selection rounds, except the BG505-sCD4 selection was performed using M48U1 in place of sCD4 (i.e., BG505-M48U1) due to its higher affinity to gp120 ($K_{D, M48U1} = 0.0154$ nM vs $K_{D, CD4} = 7.68$ nM) (30). Binders to BG505-M48U1 were collected (population 1). A third round of depletion and selection was performed using BG505-sCD4 for selection; selected

binders were collected (population 2). Finally, cells in populations 1 and 2 were sorted with and without 21c scFv, a CD4i antibody, to deplete potential off-target CD4i binders in each population. Binders in populations 3 (population 1 without 21c), 4 (population 1 with 21c), and 5 (population 2 with 21c) were deep sequenced, and ten candidates were selected for further characterization.

4.3.3 Sequence analysis and binding assays reveal five candidates bind exclusively to open Env

Enriched candidates from both in-house and Protabit selections showed high sequence diversity across CDR3 regions, consistent of successful library screening rather than clonal expansion artifacts (31). Analysis of in-house candidates revealed CDR3 lengths ranging from 8 to 20 residues, with considerable amino acid variation at each position (**Figure 5a**). CDR1 and CDR2 were fixed in length as expected (9 and 14 residues, respectively), whereas CDR3 length varied substantially. CDR3 is typically the longest loop and is usually directly responsible for antigen recognition. In our case, we expected longer loops to reach further into the caldera, and sequencing revealed that eight of our 31 candidates had the maximum CDR3 length of 20 residues (28). Protabit candidates similarly showed encouraging diversity profiles despite originating from a more constrained library design (**Figure 5b**). Their synthetic V_{HH} library was fundamentally different from our in-house approach, with all loops fixed in length and the CDR3 loop constrained to a relatively short 10-mer length. Among these ten CDR3 residues, seven positions were varied in their library design. Despite this constraint, the ten candidates they provided (P1 through P10) displayed substantial sequence variation across these variable positions, with relatively few positions sharing the same amino acid across candidates. This diversity was particularly notable given their streamlined selection protocol focused exclusively on BG505 SOSIP, suggesting that multiple distinct binding solutions to the open Env surface had been captured.

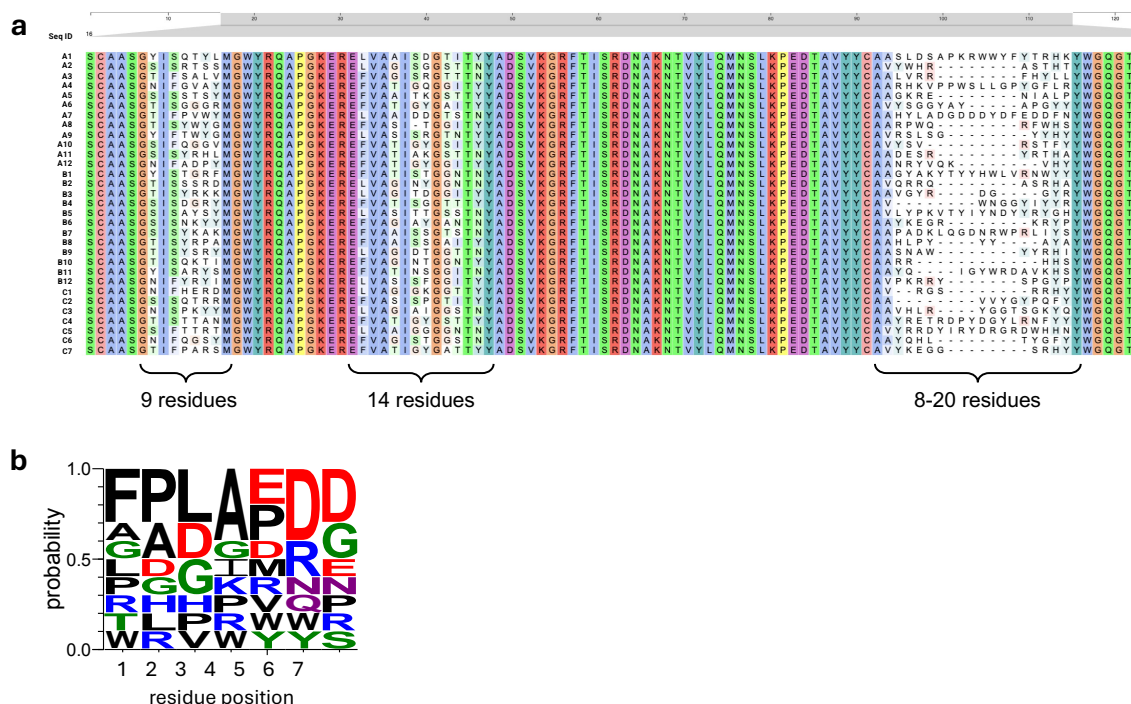


Figure 5. Sequence diversity for candidates isolated in-house and by Protabit.

(a) Sequence diversity across CDR1, CDR2, and CDR3 loops in in-house candidates.

(b) Sequence diversity across 7 residues in CDR3 loop for Protabit candidates.

To evaluate binding profiles for all 41 candidates (31 in-house and 10 from Protabit), V_{HH} -Fc constructs were expressed in 2-mL mammalian cultures using a 96 deep well plate format. Initial binding assessment was performed by ELISA using culture supernatants against plates coated with DS-BG505 (closed Env) or BG505-sCD4 (open Env) (**Figure 6a**). Potential candidates should not bind closed Env but should bind open Env. This small-scale screening revealed that none of the 31 in-house candidates bound exclusively to open Env. The in-house candidates included binders to both open and closed antigens, binders exclusive to closed antigens, and binders to neither, indicative of failed selections. However, two Protabit candidates, P4 and P5, showed clear preferential binding to BG505-sCD4 over DS-BG505.

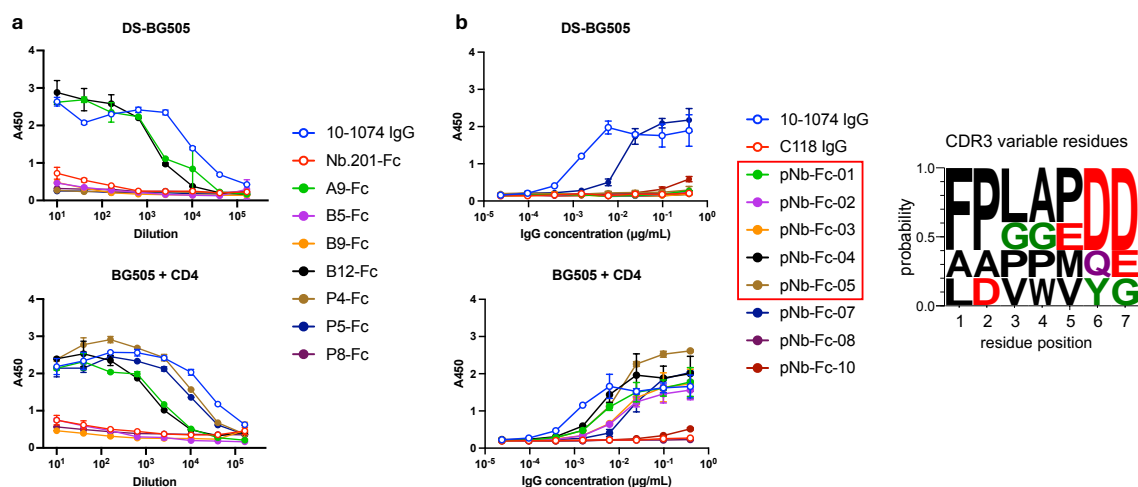


Figure 6. ELISA binding profiles for enriched in-house and Protabit candidates. 10-1074 IgG was used as a positive control (blue curves) and either Nb.201-Fc or C118 IgG was used as a negative control (red curves).

(a) Small-scale supernatant ELISA binding profiles for select candidates in both in-house and Protabit efforts.

(b) Purified protein ELISA binding profiles for all Protabit candidates that expressed well. Variable CDR3 residues for five candidates (boxed in red) shown on the right.

To validate these initial findings and assess additional candidates, all ten V_{HH}-Fc constructs were scaled up for purification. Of the ten, eight expressed successfully including P4 and P5. These eight were purified by Protein A affinity chromatography followed by SEC and reassessed by ELISA using purified protein rather than supernatant (**Figure 6b**). This analysis revealed that three additional candidates, P1-P3, also bound exclusively to open Env, bringing the total number of candidates up to five. Analysis of the CDR3 sequences for these five successful candidates (P1-P5) revealed that certain residues were shared at specific positions, with none carrying the basic residues (Lys, Arg, or His) that had been present across the broader P1-P10 set (**Figure 6b, right**). Overall, the CDR3 loop narrowed in on a select few residues in positions 1 (Phe), 2 (Pro), 6 (Asp), and 7 (Asp).

We performed additional binding assays to determine the epitope, which immediately ruled out caldera targeting. ELISA against sCD4 alone revealed that all five candidates bound

strongly to CD4, despite prior depletion steps using DS-BG505-sCD4 (**Figure 7a**). This result immediately indicated that the epitope was not the caldera because Env-bound CD4 is positioned too far from that region to allow any meaningful overlap. This suggested that sCD4 forms part of the candidates' epitope footprint.

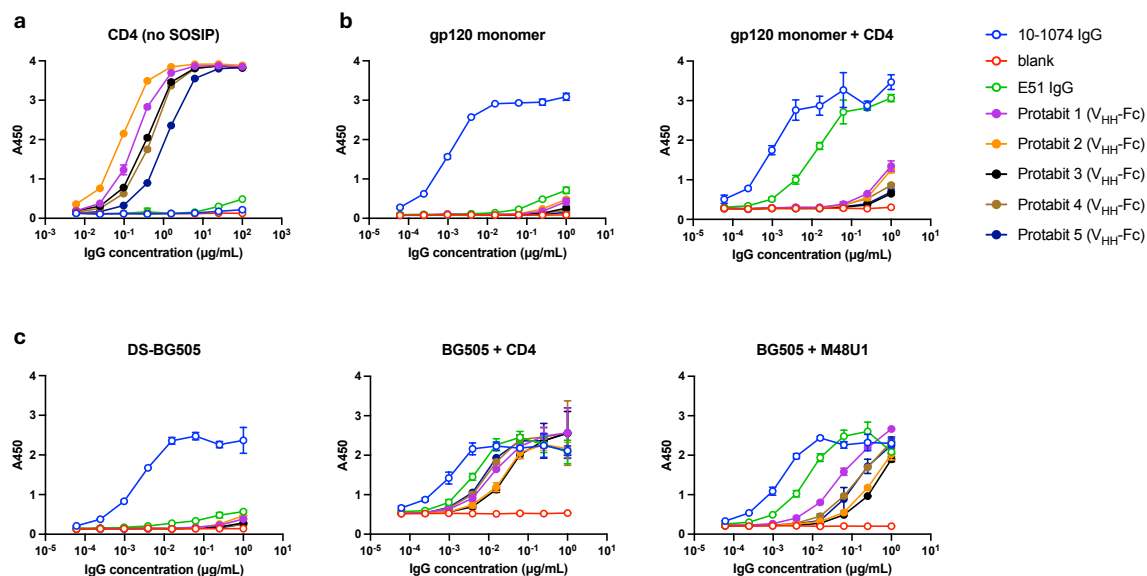


Figure 7. ELISA binding profiles for candidates against (a) sCD4 alone, (b) gp120 monomer with and without sCD4, and (c, bottom row) closed and open Env, with sCD4 or M48U1.

To assess whether the candidates might be CD4i antibodies that bind to the known coreceptor-binding site at the protomer apex, we tested binding against monomeric gp120 both with and without sCD4. Our controls performed as expected, with E51 (a known CD4i antibody) showed preferential binding to gp120 in the presence of sCD4, confirming access to the CD4i epitope (27). However, none of our five candidates showed substantial binding to gp120 monomer, with or without sCD4. This ruled out recognition of the classical CD4i epitope and supported the idea that the candidates recognize a quaternary epitope overlapping sCD4 and BG505 that requires the antigen to be trimeric.

Finally, we tested whether the candidates could bind to open Env in the absence of sCD4 using a CD4-mimetic, M48U1, that had been incorporated into Protabit's FACS workflow.

ELISAs comparing BG505-sCD4 and BG505-M48U1, both open Env, showed that all candidates maintained binding to M48U1-bound trimers, confirming that their binding was exclusive to the open state regardless of whether sCD4 or M48U1 was used to trigger conformational changes (**Figure 7c**). However, binding to BG505-M48U1 showed noticeably weaker signal intensity compared to BG505-sCD4, with modest rightward shifts in the binding curves indicative of reduced apparent affinity. This suggested that while the candidates could recognize the quaternary epitope in M48U1-bound open trimers, the epitope was partially dependent on contacts with bound sCD4. The preservation of binding in the absence of sCD4 supported our quaternary epitope theory, as binding that was completely absent on CD4-bound gp120 monomers was restored in M48U1-bound trimers.

4.3.4 Structural studies were hindered by limitations in sample preparation

Given the selective binding profiles of candidates P1-P5 to open Env, we sought to determine their binding sites through electron microscopy. However, structural characterization of Protabit candidates, in both V_{HH} and V_{HH} -Fc formats, was limited primarily by challenges in sample preparation arising from their weak binding affinities. Although binding kinetics could not be determined accurately because experiments were performed using bivalent V_{HH} -Fc constructs, ELISA data clearly indicated weaker interactions compared to control antibodies (**Figure 7c**). This weak binding became the central limitation for both negative-stain EM (nsEM) and cryogenic EM (cryo-EM) sample preparation.

Structural studies of protein complexes typically involve incubating the target with excess ligand and then purifying the assembled complex by SEC to remove any unbound components. However, for weakly bound ligands, SEC can drive dissociation because the non-equilibrium conditions across the column continuously separate free ligand from the complex, thereby favoring ligand loss. This phenomenon is well documented and was consistent with our observations (32). In our case, purification of BG505-sCD4- V_{HH} samples by standard SEC resulted in dissociation of the Protabit candidates from the complex, such that purified material no longer retained clear evidence of V_{HH} occupancy. This dissociation was evident in cryo-EM attempts (**Figure 8a**). Cryo-EM reconstruction of purified BG505-

sCD4-P5 V_{HH} complexes revealed BG505 particles coupled only with sCD4, with the P5 V_{HH} notably absent from the density map (**Figure 8a**). Attempts to image unpurified samples (to limit dissociation) were also unsuccessful as grids were too heterogeneous for particle classification, even after reducing the amount of excess ligand in the assembly reactions from the previously used 1:6:6 BG505:sCD4: V_{HH} molar ratio to 1:3.5:3.5. Similar heterogeneity was observed by nsEM, where only purified samples yielded clearly identifiable trimer particles (**Figure 8b**). Despite not obtaining structures due to low resolution, these purified complexes were likely devoid of bound V_{HH} due to dissociation during purification.

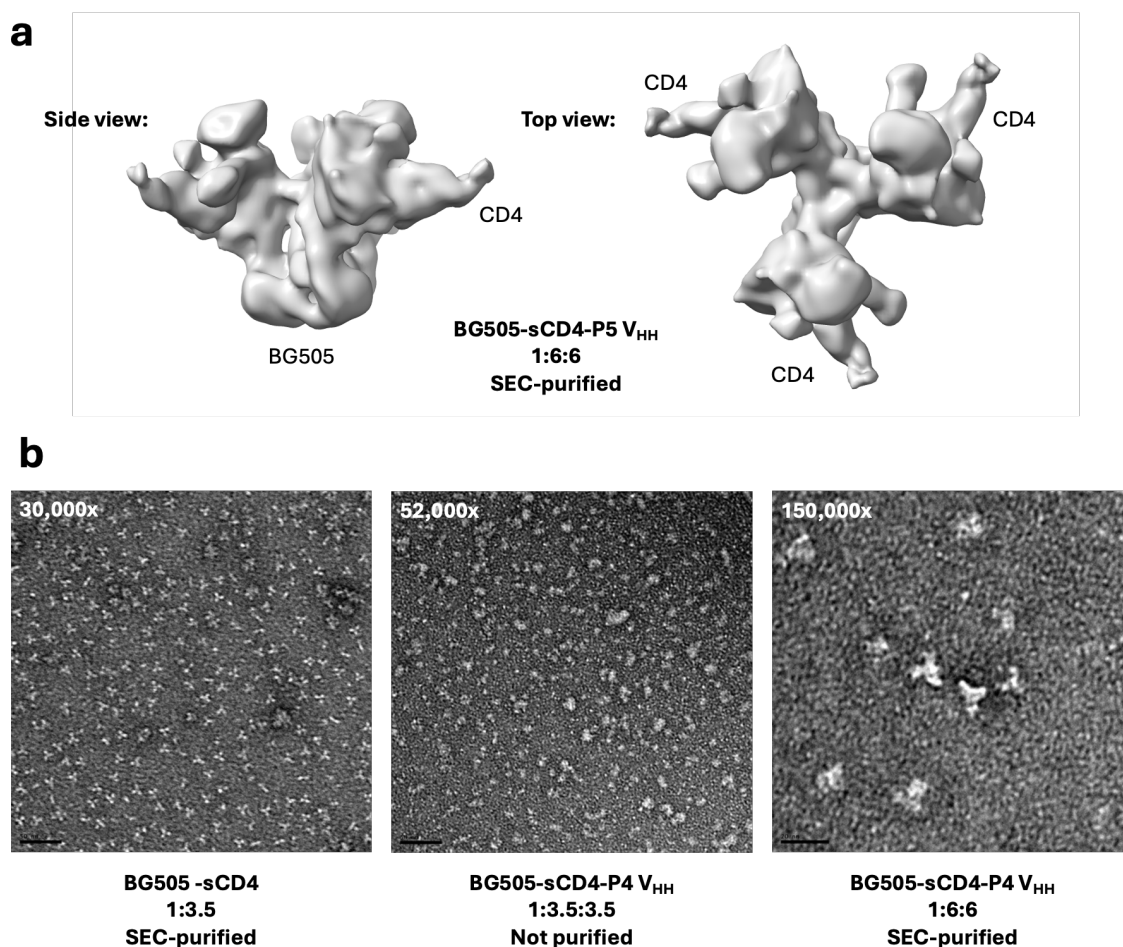


Figure 8. Attempts at structural characterization by cryo-EM and nsEM. Images courtesy of Edem Gavor.

(a) Cryo-EM map of BG505 complexed with sCD4. P5 V_{HH} (monomer) is notably absent. Similar results for P5 V_{HH}-Fc.

(b) nsEM images of BG505-sCD4 (left), unpurified BG505-sCD4-P4 V_{HH} (center), and purified BG505-sCD4-P4 V_{HH} (right).

4.3.5 Proposed epitope model based on binding constraints

By pulling together insights from the candidate binding profiles in Figure 7, we were able to define the constraints for the epitope recognized by candidates P1-P5. The data indicate that the epitope is exclusive to open Env, includes contacts with CD4, requires a quaternary trimeric antigen, is distinct from the well-characterized CD4i epitope, and remains accessible in the M48U1-bound open Env. Based on these constraints, we propose that the candidates recognize a composite interface formed between CD4 and the open Env trimer, likely positioned at or near the junction between protomers where CD4-induced conformational changes create a novel binding surface (**Figure 9**).

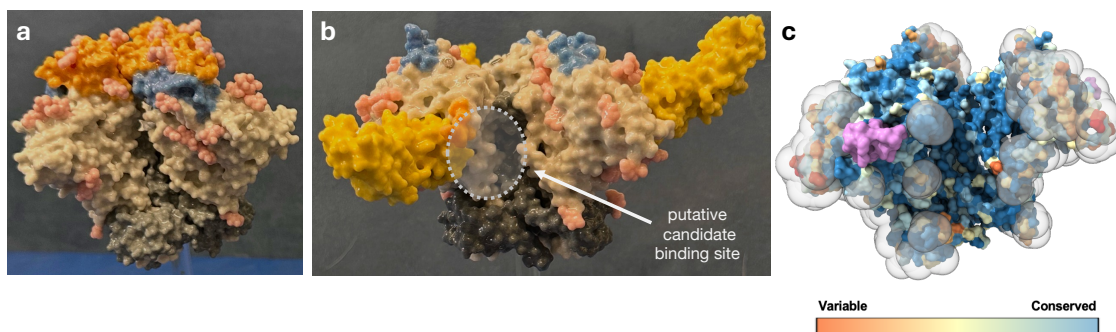


Figure 9. Proposed epitope model showing putative binding site at CD4-Env interface accessible only in the open conformation.

(a) Model of closed BG505 (PDB: 8fyi). Glycans in pink.

(b) Model of sCD4-bound open BG505 (PDB: 6u0l) with putative binding site.

(c) Model of M48U1-bound open BG505 (PDB: 7lok) with sequence conservation mapping. Glycans represented as clear regions.

Comparison of closed and CD4-bound open Env suggests that this epitope emerges from quaternary rearrangements that accompany Env opening (14, 15). The candidates' inability

to bind monomeric gp120, even in the presence of sCD4, indicates that the epitope requires contributions from multiple protomers or from conformational changes that are only possible in the context of trimers. The preservation of binding to M48U1-bound open trimers, albeit with reduced apparent affinity, suggests that while sCD4 provides optimal contacts, the underlying quaternary surface is maintained when opening is triggered by the CD4-mimetic.

4.4 Discussion

4.4.1 Simplified screens isolated binders exclusive to open Env

Our V_{HH} screening campaign yielded important insights despite not achieving its primary goal of isolating caldera-specific nanobodies. The identification of five candidates that selectively bind to a novel quaternary epitope on open Env demonstrates the effectiveness of systematic depletion and enrichment strategies for epitope discovery. While these candidates ultimately bound a CD4-overlapping interface rather than the caldera, their consistent selection across multiple enrichment rounds and their conformational specificity validate the underlying screening framework.

The success of Protabit's streamlined approach relative to the in-house multi-antigen workflow provides a clear lesson for future discovery efforts. Our initial MACS screening strategy was likely hindered by the use of an overly complex workflow built around antigens with suboptimal stability and behavior, particularly ZM197M and possibly B41. In retrospect, the lower-quality antigens likely produced poorer screening outcomes for our in-house efforts and enriched for unintended populations before even moving forward with FACS-based screening. Therefore, it is unsurprising that Protabit's streamlined protocol using only the well-characterized and best behaving BG505 SOSIP yielded candidates consistent with their screen design. By minimizing antigen-related variability, their approach enabled cleaner and more efficient enrichment.

Building off this success, future screens will incorporate M48U1 exclusively in place of sCD4. M48U1 exhibits significantly stronger binding affinity for gp120 than sCD4 ($K_{D,M48U1} = 0.0154$ nM vs. $K_{D,sCD4} = 7.68$ nM), making open trimers more stable during screening

conditions as it's less likely for Env to revert to its closed state via M48U1 disassociation (30). Additionally, the use of M48U1 eliminates the off-target enrichment to sCD4 observed in our campaign, thereby eliminating the need for complex depletion strategies. As M48U1 was already used in initial screens, eliminating sCD4 further simplifies the protocol. In downstream assays, binding profiles for sCD4-bound and M48U1-bound open Env should be near identical, providing another sanity check during ELISA-based characterizations.

4.4.2 Therapeutic potential and accessibility of the novel epitope

The identification of a novel quaternary epitope at the CD4-Env interface raises several important questions regarding its potential as a therapeutic target. Conservation analysis across HIV-1 strains indicates that this epitope is indeed well-conserved, as demonstrated by the structural mapping of sequence conservation onto the proposed binding region in Figure 9. This conservation likely reflects functional constraints imposed by CD4-induced conformational changes necessary for viral entry. While encouraging as a potential therapeutic target, caution must be taken as the candidates are susceptible to off-target effects due to them being able to bind CD4 in the absence of Env. It might be worth assessing whether our candidates are neutralizing when linked to sCD4, consistent with previous pseudovirus neutralization experiments using 17b scFv which found only linked constructs showed some neutralizing activity (33, 34). This might not be feasible in our case, however, as we expect candidate binding to linked sCD4 to result in aggregation.

The accessibility of this epitope to conventional antibodies remains an important unresolved question as well. All five candidates from our screen were single-domain V_{HH} nanobodies with relatively short CDR3 loops (ten residues), despite the convention that longer loops would be advantageous for accessing cryptic or recessed epitopes on the Env surface (35). The success of compact nanobodies with constrained CDR3 regions could indicate either that the epitope is more structurally accessible than initially assumed, or alternatively that these particular candidates are engaging BG505-specific glycans arrangements rather than a cryptic protein epitope. While structural analysis of the BG505 structure indicates the region of interest likely lacks *N*-linked glycans, comprehensive structural characterization of these

complexes is essential to distinguish between binding to a quaternary protein interface versus recognition of a strain-specific glycan pattern that happens to be CD4-dependent and is difficult to visualize on structures. Should the epitope in fact be on the protein surface, the question that follows is whether larger antibodies can access the same region given the potential steric occlusion by nearby glycans. This is especially true when considering the V1V2 loops, which are disordered in open Env and can carry *N*-linked glycans capable of limiting binder access (14, 36). If this is the case, this may explain in part why antibodies against this epitope have yet to be isolated and characterized. Nonetheless, while the epitope is novel, binding to CD4 ultimately limits its practical use as a therapeutic target.

4.4.3 Why CD4i antibodies have been isolated but anti-caldera antibodies have not

CD4i antibodies such as 17b, E51, and 21c have been characterized since the early 1990s and continue to expand as a class, while no antibody recognizing the caldera has been reported despite its high conservation and similar accessibility conditions in open Env (25–27). Although both epitope classes are unmasked by CD4 engagement, highly conserved, and absent in closed trimers, several key factors help explain why one has been extensively characterized while the other remains elusive.

CD4i epitopes lie on the outward-facing surface of individual gp120 monomers, including the bridging sheet and the base of V3, and can be fully reconstituted on monomeric gp120-CD4 complexes without contributions from neighboring protomers (11). The caldera, by contrast, is a quaternary surface that requires the simultaneous outward rotation of all three gp120 protomers and does not exist on a monomeric gp120, closed trimers, or any partial assembly (14, 15, 17). The minimal antigenic unit for caldera recognition is therefore the fully open trimer. The same constraint may extend to the quaternary epitope identified in this study, which has likewise not been previously characterized. However, monomeric gp120 served as the standard HIV-1 immunogen and discovery bait for decades, ever since its co-crystal structure with CD4 and 17b Fab was solved (11). CD4i antibodies isolated during that time were recovered using this monomeric reagent. No analogous trimeric reagent existed for the caldera until stable soluble Env trimers became available beginning with

BG505 SOSIP.664 in 2013 (12). Cryo-EM structures that actually defined and characterized the caldera only began appearing in 2016 and 2017, meaning the caldera has been screenable for about a decade compared to over three decades for CD4i sites (14, 15).

Antigen exposure also plays a role in the lack of isolated anti-caldera antibodies. Productive opening of Env at the cell surface during infection is brief, irreversible, and rapidly succeeded by membrane fusion, making the open trimer a kinetic intermediate rather than a steady-state population (37, 38). CD4i antibodies can be elicited by shed gp120 or by gp120 in transient association with cellular CD4, neither of which requires the trimer to be intact (39–41). The caldera, however, requires that the entire trimer remain assembled and open long enough to engage rare germline B cell precursors, a configuration unlikely to persist during infection. The same may apply to the quaternary epitope identified herein if its presentation depends on the same short-lived open-state intermediate. Consequently, natural infection readily elicits non-neutralizing CD4i responses but mounts essentially no detectable response against the caldera.

Despite these challenges, caldera targeting appears feasible in principle. The surface has been visualized by cryo-EM, is solvent-accessible in the open trimer, and is largely free of glycan shielding (14, 15, 17). Small protein scaffolds, particularly V_{HH} nanobodies and *de novo* designed miniproteins, have repeatedly demonstrated the ability to access similarly recessed pockets on other challenging targets (20, 42). In recent years, computational binder design has matured rapidly, with tools like RFdiffusion, BoltzGen, and ProteinMPNN now routinely producing picomolar binders against difficult targets without requiring any natural template or starting point, so potential caldera binders may be generated and enriched in our improving screening workflow in the coming years (43–46).

4.4.4 Future directions

Ongoing efforts by Protabit are already began implementing M48U1-exclusive screening strategies combined with BG505-3fill trimers with three engineered *N*-linked glycans to block (or “fill”) an immunogenic hole in the BG505 glycan shield (47). Additionally,

computational in-house efforts utilizing BoltzGen to generate binders are underway. Should an anti-caldera binder be isolated from either effort, we would implement the same characterization workflow described herein with the hope of arriving at structures of candidates bound to the Env caldera. Once validated, we can also co-develop a bi-functional molecule by fusing the caldera binder to an engineered CD4 that is specific to gp120 (48). This construct would combine CD4-mediated Env opening with caldera-specific engagement, allowing caldera binders to bind on free virions before they infect healthy cells. Finally, our methodological advances developed for HIV-1 caldera targeting can be translated to other challenging viral targets, especially those concerning transient viral epitopes such as in influenza HA stem regions and RSV F prefusion sites, where similar challenges with epitope accessibility and conformational specificity arise (49, 50).

4.5 Materials and Methods

4.5.1 Protein expression and purification

Mammalian transfections for all constructs in this study were performed by the Protein Expression Center (PEC) using Expi293 cells. SOSIP trimers were purified over 5-mL PGT145 affinity columns with PBS (20 mM sodium phosphate, 150 mM NaCl, pH 7.4) as running buffer and eluted with 3M MgCl₂. Size-exclusion chromatography (SEC) was then performed over Superose 6 Increase 10/300 GL equilibrated with PBS. sCD4 constructs were His-tagged and purified with Ni-NTA affinity columns (PBS pH 7.4 running buffer, 100 mM imidazole pH 7.4 elution buffer) before purifying over SEC with Superdex 200 Increase 10/300 GL equilibrated with PBS (pH 7.4). V_{HH}-Fc candidates were purified by Protein A affinity columns (PBS pH 7.4 running buffer, 100 mM HCl-glycine pH 3 elution buffer) before purifying over SEC with Superdex 200 Increase 10/300 GL equilibrated with PBS (pH 7.4). Molecular weights for all purified samples were confirmed via SDS-PAGE.

4.5.2 SOSIP labeling with amine chemistry

SOSIPs were randomly AF647-labeled using Alexa Fluor 647 NHS ester (succinimidyl ester) (Thermo Fisher Scientific). SOSIP at 5 μ M was mixed with 100 μ M Alexa Fluor 647 NHS ester (succinimidyl ester) (20X excess dye) in PBS, pH 7.4 and placed on a rocking

platform at 4°C overnight. Zeba spin desalting columns (7 kDa MWCO) were then used to remove excess dye. SOSIP preps were split into 50-µg aliquots before flash freezing and storing in -80°C. The same protocol was used for randomly biotinylating SOSIPs, except with EZ-Link NHS-PEG4-Biotin kit (Thermo Fisher Scientific) in place of AF647. Similarly, zeba spin columns (7 kDa MWCO) were used to remove excess biotin before flash freezing and storing 50-µg aliquots.

4.5.3 Yeast display library screening

A synthetic V_{HH} nanobody yeast display library developed in the Kruse laboratory was screened by iterative magnetic-activated cell sorting (MACS) followed by fluorescence-activated cell sorting (FACS), adapted from the yeast-display nanobody discovery workflow of McMahon et al (28). Yeast cultures were expanded between enrichment rounds. MACS enrichment consisted of three rounds. In the first round, negative selection was performed sequentially with DS-BG505, DS-B41, and DS-ZM197M to deplete binders to closed Env. For each depletion step, induced nanobody-displaying yeast were incubated with AF647-labeled antigen, washed to remove excess unbound probe, incubated with anti-AF647 magnetic beads, washed again, and applied by gravity to an LD depletion column at 4°C; cells retained on the column were discarded and the flowthrough was carried forward. In the second round, negative selection was repeated sequentially with DS-BG505, DS-B41, and DS-ZM197M pre-complexed with sCD4 at a 1:6 SOSIP:sCD4 molar ratio to deplete binders to sCD4 and to epitopes formed at or near the SOSIP-sCD4 interface. Following these depletion steps, positive selection was performed with BG505 pre-complexed with sCD4 at a 1:6 BG505:sCD4 molar ratio using an LS selection column, and retained cells were eluted and expanded. In the third MACS round, positive selections were performed sequentially with BG505-sCD4, B41-sCD4, and ZM197M-sCD4, each pre-complexed at a 1:6 SOSIP:sCD4 molar ratio, using LS columns to further enrich for open-Env binders. The final enriched population was expanded before FACS.

FACS enrichment was then performed over three rounds using AF647-labeled SOSIP-sCD4 probes and AF488-labeled anti-HA antibody to monitor display of HA-tagged nanobodies.

In each round, yeast were gated to distinguish non-displaying cells, displaying non-binders, and double-positive displaying binders. The double-positive population was collected for expansion. Each FACS round was split into two parallel groups. In group A, cells were screened with SOSIP-sCD4 complexes alone. In group B, antigens were pre-incubated with both sCD4 and 21c scFv to block the 21c-defined CD4-induced epitope and thereby reduce recovery of binders targeting that region. The first, second, and third FACS rounds used BG505-sCD4, B41-sCD4, and ZM197M-sCD4, respectively, with matched 21c-containing conditions in the parallel group. Cell pellets collected after MACS and after each FACS round were subjected to deep sequencing, and candidate V_{HH} s were selected from enriched sequence populations for downstream characterization.

4.5.4 Deep sequencing and sequence analysis

Deep sequencing of enriched V_{HH} populations was performed by Protabit using a proprietary nanopore sequencing approach adapted from the R2C2 method for improved accuracy (51). Briefly, target V_{HH} regions were PCR-amplified using campaign-specific primers designed to capture the full variable domain. Amplicons were then circularized and subjected to rolling circle amplification (RCA) to generate concatemers containing multiple copies of each unique sequence. The RCA products were debranched and size-selected by gel electrophoresis to balance the number of sequence repeats per molecule with total read count. Size-selected products were processed using a standard Oxford Nanopore sequencing workflow.

Raw nanopore reads were processed using the C3poa package to generate highly accurate consensus sequences from the concatemeric reads, leveraging the multiple copies of each sequence within individual molecules to correct sequencing errors. Campaign-specific downstream processing was applied to identify unique V_{HH} variants and quantify their relative abundances across enrichment rounds.

For sequence analysis, V_{HH} candidates with high sequence counts and consistent enrichment patterns across multiple rounds were selected for further characterization. Multiple sequence

alignment of variable domain sequences was performed using Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>). Sequence logo plots were generated using WebLogo 3 (<http://weblogo.threeplusone.com/>) for the Protabit candidates, which had uniform CDR3 lengths suitable for positional analysis. In-house candidates, which exhibited variable CDR3 lengths ranging from 8 to 20 residues, were analyzed for length distribution and diversity patterns but were not included in logo plot analysis due to length heterogeneity.

4.5.5 Enzyme-linked immunosorbent assay (ELISA)

Streptavidin-coated 96-well plates (Thermo Fisher Scientific) were coated with 50 μ L per well of 5 μ g/mL randomly biotinylated BG505 or DS-BG505 and incubated for 1 hour at room temperature. Plates were then washed with TBS-T to remove excess unbound antigen. For assays evaluating binding to open Env, BG505-coated wells were incubated with sCD4 at 100 μ g/mL for at least 2 hours at room temperature, while DS-BG505-coated wells received an equivalent volume of TBS-T for consistency. The sCD4-containing solution was not removed before addition of candidates in order to minimize sCD4 dissociation from BG505. Therefore, after addition of an equal volume of candidate, the final sCD4 concentration in BG505 wells was 50 μ g/mL. Candidates were added directly to wells in duplicate as serial dilutions and incubated for 2 hours at room temperature. Purified proteins were tested as four-fold serial dilutions beginning at 200 μ g/mL, yielding an initial in-well concentration of 100 μ g/mL, whereas supernatants were tested as four-fold serial dilutions beginning at 5X, yielding an initial in-well dilution of 10X. Plates were then washed twice with TBS-T. Bound candidates were detected with 50 μ L per well of HRP-conjugated anti-human Fc secondary antibody (Abcam) at a 1:10,000 dilution for 1 hour at room temperature. After two additional washes with TBS-T, 50 μ L of Ultra TMB-ELISA substrate solution (Thermo Fisher Scientific) was added to each well and incubated for 10 min before quenching with 50 μ L of 1.0 N HCl. Absorbance was measured at 450 nm.

4.5.6 Cryo-EM

Cryo-electron microscopy was performed as described previously to determine high-resolution structures of V_{HH} -Env complexes (17). Briefly, V_{HH} candidates (in V_{HH} or V_{HH} -

Fc formats) were incubated with open BG505-sCD4 Env complexes at a 1:3.5:3.5 molar ratio and overnight at 4°C. Multiple purification strategies were tested: unpurified samples were used to preserve weak interactions, while SEC with PBS pH 7.4 running buffer was performed using analytical Superose 6 Increase 3.2/300 columns at flow rates of 0.01-0.02 mL/min and preparative Superose 6 Increase 10/300 GL columns to remove excess unbound components. Complex formation and stoichiometry were validated by mass photometry before grid preparation. Aliquots of 3 µL were applied to grids, blotted for 3 seconds with a blot force of 1 using a Vitrobot Mark IV, and plunge-frozen in liquid ethane. Grids were imaged using a Titan Arctica transmission electron microscope with a 6k x 4k Gatan K3 direct electron detector and running SerialEM package for automated data collection. Particle picking, 2D classification, and 3D reconstruction were performed using RELION.

4.5.7 Negative-stain EM

Negative-stain electron microscopy was performed as described previously to assess V_{HH} -Fc binding to BG505-sCD4 complexes (14). Briefly, candidate V_{HH} or V_{HH} -Fc was complexed with BG505 after overnight incubation with sCD4 at 4°C using a 1:6:6 molar ratio of BG505:sCD4: V_{HH} (-Fc) for complexes undergoing SEC purification, or a 1:3.5:3.5 molar ratio for samples requiring unpurified preparation. SEC was performed using a Superose 6 Increase 10/300 GL column equilibrated in PBS. For unpurified samples, excess unbound components were diluted to minimize background interference during grid preparation. Samples were applied to glow-discharged carbon-coated copper grids, blotted, and stained with 2% (w/v) uranyl acetate. Grids were imaged using a Tecnai Spirit transmission electron microscope operated at 120 kV. Images were collected at magnifications ranging from 52,000x to 150,000x to visualize individual Env trimers and assess complex formation.

4.6 References

1. L. Corey, P. B. Gilbert, M. Juraska, D. C. Montefiori, L. Morris, S. T. Karuna, S. Edupuganti, N. M. Mgodhi, A. C. deCamp, E. Rudnicki, Y. Huang, P. Gonzales, R. Cabello, C. Orrell, J. R. Lama, F. Laher, E. M. Lazarus, J. Sanchez, I. Frank, J. Hinojosa, M. E. Sobieszczyk, K. E. Marshall, P. G. Mukwekwerere, J. Makhema, L. R. Baden, J. I. Mullins, C. Williamson, J. Hural, M. J. McElrath, C. Bentley, S. Takuva, M. M. G. Lorenzo, D. N. Burns, N. Espy, A. K. Randhawa, N. Kochar, E. Piwowar-Manning, D. J. Donnell, N. Sista, P. Andrew, J. G. Kublin, G. Gray, J. E. Ledgerwood, J. R. Mascola, M. S. Cohen, Two Randomized Trials of Neutralizing Antibodies to Prevent HIV-1 Acquisition. *N. Engl. J. Med.* **384**, 1003–1014 (2021).
2. M. Caskey, F. Klein, J. C. C. Lorenzi, M. S. Seaman, A. P. West, N. Buckley, G. Kremer, L. Nogueira, M. Braunschweig, J. F. Scheid, J. A. Horwitz, I. Shimeliovich, S. Ben-Avraham, M. Witmer-Pack, M. Platten, C. Lehmann, L. A. Burke, T. Hawthorne, R. J. Gorelick, B. D. Walker, T. Keler, R. M. Gulick, G. Fätkenheuer, S. J. Schlesinger, M. C. Nussenzweig, Viraemia suppressed in HIV-1-infected humans by broadly neutralizing antibody 3BNC117. *Nature* **522**, 487–491 (2015).
3. M. Caskey, T. Schoofs, H. Gruell, A. Settler, T. Karagounis, E. F. Kreider, B. Murrell, N. Pfeifer, L. Nogueira, T. Y. Oliveira, G. H. Learn, Y. Z. Cohen, C. Lehmann, D. Gillor, I. Shimeliovich, C. Unson-O'Brien, D. Weiland, A. Robles, T. Kümmerle, C. Wyen, R. Levin, M. Witmer-Pack, K. Eren, C. Ignacio, S. Kiss, A. P. West, H. Mouquet, B. S. Zingman, R. M. Gulick, T. Keler, P. J. Bjorkman, M. S. Seaman, B. H. Hahn, G. Fätkenheuer, S. J. Schlesinger, M. C. Nussenzweig, F. Klein, Antibody 10-1074 suppresses viremia in HIV-1-infected individuals. *Nat. Med.* **23**, 185–191 (2017).
4. P. Mendoza, H. Gruell, L. Nogueira, J. A. Pai, A. L. Butler, K. Millard, C. Lehmann, I. Suárez, T. Y. Oliveira, J. C. C. Lorenzi, Y. Z. Cohen, C. Wyen, T. Kümmerle, T. Karagounis, C.-L. Lu, L. Handl, C. Unson-O'Brien, R. Patel, C. Ruping, M. Schlotz, M. Witmer-Pack, I. Shimeliovich, G. Kremer, E. Thomas, K. E. Seaton, J. Horowitz, A. P. West, P. J. Bjorkman, G. D. Tomaras, R. M. Gulick, N. Pfeifer, G. Fätkenheuer, M. S. Seaman, F. Klein, M. Caskey, M. C. Nussenzweig, Combination therapy with anti-HIV-1 antibodies maintains viral suppression. *Nature* **561**, 479–484 (2018).
5. J. F. Scheid, J. A. Horwitz, Y. Bar-On, E. F. Kreider, C.-L. Lu, J. C. C. Lorenzi, A. Feldmann, M. Braunschweig, L. Nogueira, T. Oliveira, I. Shimeliovich, R. Patel, L. Burke, Y. Z. Cohen, S. Hadrigan, A. Settler, M. Witmer-Pack, J. West, B. Juelg, T. Keler, T. Hawthorne, B. Zingman, R. M. Gulick, N. Pfeifer, G. H. Learn, M. S. Seaman, P. J. Bjorkman, F. Klein, S. J. Schlesinger, B. D. Walker, B. H. Hahn, M. C.

- Nussenzweig, M. Caskey, HIV-1 antibody 3BNC117 suppresses viral rebound in humans during treatment interruption. *Nature* **535**, 556–560 (2016).
6. C. Ding, D. Patel, Y. Ma, J. F. S. Mann, J. Wu, Y. Gao, Employing Broadly Neutralizing Antibodies as a Human Immunodeficiency Virus Prophylactic & Therapeutic Application. *Front. Immunol.* **12** (2021).
 7. M. D. Miner, L. Corey, D. Montefiori, Broadly neutralizing monoclonal antibodies for HIV prevention. *J. Int. AIDS Soc.* **24**, e25829 (2021).
 8. S. A. Griffith, L. E. McCoy, To bnAb or Not to bnAb: Defining Broadly Neutralising Antibodies Against HIV-1. *Front. Immunol.* **12** (2021).
 9. H. Choe, M. Farzan, Y. Sun, N. Sullivan, B. Rollins, P. D. Ponath, L. Wu, C. R. Mackay, G. LaRosa, W. Newman, N. Gerard, C. Gerard, J. Sodroski, The β -Chemokine Receptors CCR3 and CCR5 Facilitate Infection by Primary HIV-1 Isolates. *Cell* **85**, 1135–1148 (1996).
 10. T. Dragic, V. Litwin, G. P. Allaway, S. R. Martin, Y. Huang, K. A. Nagashima, C. Cayanan, P. J. Maddon, R. A. Koup, J. P. Moore, W. A. Paxton, HIV-1 entry into CD4⁺ cells is mediated by the chemokine receptor CC-CKR-5. *Nature* **381**, 667–673 (1996).
 11. P. D. Kwong, R. Wyatt, J. Robinson, R. W. Sweet, J. Sodroski, W. A. Hendrickson, Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* **393**, 648–659 (1998).
 12. R. W. Sanders, R. Derking, A. Cupo, J.-P. Julien, A. Yasmeen, N. de Val, H. J. Kim, C. Blattner, A. T. de la Peña, J. Korzun, M. Golabek, K. de los Reyes, T. J. Ketas, M. J. van Gils, C. R. King, I. A. Wilson, A. B. Ward, P. J. Klasse, J. P. Moore, A Next-Generation Cleaved, Soluble HIV-1 Env Trimer, BG505 SOSIP.664 gp140, Expresses Multiple Epitopes for Broadly Neutralizing but Not Non-Neutralizing Antibodies. *PLOS Pathog.* **9**, e1003618 (2013).
 13. A. B. Ward, I. A. Wilson, The HIV-1 envelope glycoprotein structure: nailing down a moving target. *Immunol. Rev.* **275**, 21–32 (2017).
 14. H. Wang, A. A. Cohen, R. P. Galimidi, H. B. Gristick, G. J. Jensen, P. J. Bjorkman, Cryo-EM structure of a CD4-bound open HIV-1 envelope trimer reveals structural rearrangements of the gp120 V1V2 loop. *Proc. Natl. Acad. Sci.* **113**, E7151–E7158 (2016).

15. G. Ozorowski, J. Pallesen, N. de Val, D. Lyumkis, C. A. Cottrell, J. L. Torres, J. Copps, R. L. Stanfield, A. Cupo, P. Pugach, J. P. Moore, I. A. Wilson, A. B. Ward, Open and closed structures reveal allostery and pliability in the HIV-1 envelope spike. *Nature* **547**, 360–363 (2017).
16. H. Wang, C. O. Barnes, Z. Yang, M. C. Nussenzweig, P. J. Bjorkman, Partially Open HIV-1 Envelope Structures Exhibit Conformational Changes Relevant for Coreceptor Binding and Fusion. *Cell Host Microbe* **24**, 579-592.e4 (2018).
17. Z. Yang, H. Wang, A. Z. Liu, H. B. Gristick, P. J. Bjorkman, Asymmetric opening of HIV-1 Env bound to CD4 and a coreceptor-mimicking antibody. *Nat. Struct. Mol. Biol.* **26**, 1167–1175 (2019).
18. C. A. Jette, C. O. Barnes, S. M. Kirk, B. Melillo, A. B. Smith, P. J. Bjorkman, Cryo-EM structures of HIV-1 trimer bound to CD4-mimetics BNM-III-170 and M48U1 adopt a CD4-bound open conformation. *Nat. Commun.* **12**, 1950 (2021).
19. P. Holliger, P. J. Hudson, Engineered antibody fragments and the rise of single domains. *Nat. Biotechnol.* **23**, 1126–1136 (2005).
20. S. Muyldermans, Nanobodies: Natural Single-Domain Antibodies. *Annu. Rev. Biochem.* **82**, 775–797 (2013).
21. P. Pugach, G. Ozorowski, A. Cupo, R. Ringe, A. Yasmeen, N. de Val, R. Derking, H. J. Kim, J. Korzun, M. Golabek, K. de los Reyes, T. J. Ketas, J.-P. Julien, D. R. Burton, I. A. Wilson, R. W. Sanders, P. J. Klasse, A. B. Ward, J. P. Moore, A Native-Like SOSIP.664 Trimer Based on an HIV-1 Subtype B env Gene. *J. Virol.* **89**, 3380–3395 (2015).
22. J.-P. Julien, J. H. Lee, G. Ozorowski, Y. Hua, A. Torrents de la Peña, S. W. de Taeye, T. Nieusma, A. Cupo, A. Yasmeen, M. Golabek, P. Pugach, P. J. Klasse, J. P. Moore, R. W. Sanders, A. B. Ward, I. A. Wilson, Design and structure of two HIV-1 clade C SOSIP.664 trimers that increase the arsenal of native-like Env immunogens. *Proc. Natl. Acad. Sci.* **112**, 11947–11952 (2015).
23. Y. Do Kwon, M. Pancera, P. Acharya, I. S. Georgiev, E. T. Crooks, J. Gorman, M. G. Joyce, M. Guttman, X. Ma, S. Narpala, C. Soto, D. S. Terry, Y. Yang, T. Zhou, G. Ahlsen, R. T. Bailer, M. Chambers, G.-Y. Chuang, N. A. Doria-Rose, A. Druz, M. A. Hallen, A. Harned, T. Kirys, M. K. Louder, S. O'Dell, G. Ofek, K. Osawa, M. Prabhakaran, M. Sastry, G. B. E. Stewart-Jones, J. Stuckey, P. V. Thomas, T. Tittley, C. Williams, B. Zhang, H. Zhao, Z. Zhou, B. R. Donald, L. K. Lee, S. Zolla-Pazner, U. Baxa, A. Schön, E. Freire, L. Shapiro, K. K. Lee, J. Arthos, J. B. Munro, S. C.

- Blanchard, W. Mothes, J. M. Binley, A. B. McDermott, J. R. Mascola, P. D. Kwong, Crystal structure, conformational fixation and entry-related interactions of mature ligand-free HIV-1 Env. *Nat. Struct. Mol. Biol.* **22**, 522–531 (2015).
24. S. W. de Taeye, G. Ozorowski, A. Torrents de la Peña, M. Guttman, J.-P. Julien, T. L. G. M. van den Kerkhof, J. A. Burger, L. K. Pritchard, P. Pugach, A. Yasmeen, J. Crampton, J. Hu, I. Bontjer, J. L. Torres, H. Arendt, J. DeStefano, W. C. Koff, H. Schuitemaker, D. Eggink, B. Berkhout, H. Dean, C. LaBranche, S. Crotty, M. Crispin, D. C. Montefiori, P. J. Klasse, K. K. Lee, J. P. Moore, I. A. Wilson, A. B. Ward, R. W. Sanders, Immunogenicity of Stabilized HIV-1 Envelope Trimers with Reduced Exposure of Non-neutralizing Epitopes. *Cell* **163**, 1702–1715 (2015).
 25. M. Thali, J. P. Moore, C. Furman, M. Charles, D. D. Ho, J. Robinson, J. Sodroski, Characterization of conserved human immunodeficiency virus type 1 gp120 neutralization epitopes exposed upon gp120-CD4 binding. *J. Virol.* **67**, 3978–3988 (1993).
 26. R. Diskin, P. M. Marcovecchio, P. J. Bjorkman, Structure of a clade C HIV-1 gp120 bound to CD4 and CD4-induced antibody reveals anti-CD4 polyreactivity. *Nat. Struct. Mol. Biol.* **17**, 608–613 (2010).
 27. S.-H. Xiang, L. Wang, M. Abreu, C.-C. Huang, P. D. Kwong, E. Rosenberg, J. E. Robinson, J. Sodroski, Epitope mapping and characterization of a novel CD4-induced human monoclonal antibody capable of neutralizing primary HIV-1 strains. *Virology* **315**, 124–134 (2003).
 28. C. McMahon, A. S. Baier, R. Pascolutti, M. Wegrecki, S. Zheng, J. X. Ong, S. C. Erlandson, D. Hilger, S. G. F. Rasmussen, A. M. Ring, A. Manglik, A. C. Kruse, Yeast surface display platform for rapid discovery of conformationally selective nanobodies. *Nat. Struct. Mol. Biol.* **25**, 289–296 (2018).
 29. Q. Liu, P. Acharya, M. A. Dolan, P. Zhang, C. Guzzo, J. Lu, A. Kwon, D. Gururani, H. Miao, T. Bylund, G.-Y. Chuang, A. Druz, T. Zhou, W. J. Rice, C. Wigge, B. Carragher, C. S. Potter, P. D. Kwong, P. Lusso, Quaternary contact in the initial interaction of CD4 with the HIV-1 envelope trimer. *Nat. Struct. Mol. Biol.* **24**, 370–378 (2017).
 30. P. Acharya, T. S. Luongo, M. K. Louder, K. McKee, Y. Yang, Y. Do Kwon, J. R. Mascola, P. Kessler, L. Martin, P. D. Kwong, Structural Basis for Highly Effective HIV-1 Neutralization by CD4-Mimetic Miniproteins Revealed by 1.5 Å Cocrystal Structure of gp120 and M48U1. *Structure* **21**, 1018–1029 (2013).

31. R. Rouet, K. J. L. Jackson, D. B. Langley, D. Christ, Next-Generation Sequencing of Antibody Display Repertoires. *Front. Immunol.* **9** (2018).
32. K. F. Blom, B. S. Larsen, C. N. McEwen, Determining Affinity-Selected Ligands and Estimating Binding Affinities by Online Size Exclusion Chromatography/Liquid Chromatography–Mass Spectrometry. *J. Comb. Chem.* **1**, 82–90 (1999).
33. A. P. West, R. P. Galimidi, C. P. Foglesong, P. N. P. Gnanapragasam, J. S. Klein, P. J. Bjorkman, Evaluation of CD4-CD4i Antibody Architectures Yields Potent, Broadly Cross-Reactive Anti-Human Immunodeficiency Virus Reagents. *J. Virol.* **84**, 261–269 (2010).
34. L. A. Lagenaur, V. A. Villarroel, V. Bundoc, B. Dey, E. A. Berger, sCD4-17b bifunctional protein: Extremely broad and potent neutralization of HIV-1 Env pseudotyped viruses from genetically diverse primary isolates. *Retrovirology* **7**, 11 (2010).
35. M. I. J. Raybould, C. Marks, K. Krawczyk, B. Taddese, J. Nowak, A. P. Lewis, A. Bujotzek, J. Shi, C. M. Deane, Five computational developability guidelines for therapeutic antibody profiling. *Proc. Natl. Acad. Sci.* **116**, 4025–4030 (2019).
36. J. S. McLellan, M. Pancera, C. Carrico, J. Gorman, J.-P. Julien, R. Khayat, R. Louder, R. Pejchal, M. Sastry, K. Dai, S. O'Dell, N. Patel, S. Shahzad-ul-Hussan, Y. Yang, B. Zhang, T. Zhou, J. Zhu, J. C. Boyington, G.-Y. Chuang, D. Diwanji, I. Georgiev, Y. Do Kwon, D. Lee, M. K. Louder, S. Moquin, S. D. Schmidt, Z.-Y. Yang, M. Bonsignori, J. A. Crump, S. H. Kapiga, N. E. Sam, B. F. Haynes, D. R. Burton, W. C. Koff, L. M. Walker, S. Phogat, R. Wyatt, J. Orwenyo, L.-X. Wang, J. Arthos, C. A. Bewley, J. R. Mascola, G. J. Nabel, W. R. Schief, A. B. Ward, I. A. Wilson, P. D. Kwong, Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. *Nature* **480**, 336–343 (2011).
37. J. B. Munro, J. Gorman, X. Ma, Z. Zhou, J. Arthos, D. R. Burton, W. C. Koff, J. R. Courter, A. B. Smith, P. D. Kwong, S. C. Blanchard, W. Mothes, Conformational dynamics of single HIV-1 envelope trimers on the surface of native virions. *Science* **346**, 759–763 (2014).
38. X. Ma, M. Lu, J. Gorman, D. S. Terry, X. Hong, Z. Zhou, H. Zhao, R. B. Altman, J. Arthos, S. C. Blanchard, P. D. Kwong, J. B. Munro, W. Mothes, HIV-1 Env trimer opens through an asymmetric intermediate in which individual protomers adopt distinct conformations. *eLife* **7**, e34271 (2018).

39. J. P. Moore, J. A. McKeating, R. A. Weiss, Q. J. Sattentau, Dissociation of gp120 from HIV-1 Virions Induced by Soluble CD4. *Science* **250**, 1139–1142 (1990).
40. T. K. Hart, R. Kirsh, H. Ellens, R. W. Sweet, D. M. Lambert, S. R. Petteway, J. Leary, P. J. Bugelski, Binding of soluble CD4 proteins to human immunodeficiency virus type 1 and infected cells induces release of envelope glycoprotein gp120. *Proc. Natl. Acad. Sci.* **88**, 2189–2193 (1991).
41. R. Wyatt, P. D. Kwong, E. Desjardins, R. W. Sweet, J. Robinson, W. A. Hendrickson, J. G. Sodroski, The antigenic structure of the HIV gp120 envelope glycoprotein. *Nature* **393**, 705–711 (1998).
42. L. Cao, I. Goresnik, B. Coventry, J. B. Case, L. Miller, L. Kozodoy, R. E. Chen, L. Carter, A. C. Walls, Y.-J. Park, E.-M. Strauch, L. Stewart, M. S. Diamond, D. Veessler, D. Baker, De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. *Science* **370**, 426–431 (2020).
43. J. L. Watson, D. Juergens, N. R. Bennett, B. L. Trippe, J. Yim, H. E. Eisenach, W. Ahern, A. J. Borst, R. J. Ragotte, L. F. Milles, B. I. M. Wicky, N. Hanikel, S. J. Pellock, A. Courbet, W. Sheffler, J. Wang, P. Venkatesh, I. Sappington, S. V. Torres, A. Lauko, V. De Bortoli, E. Mathieu, S. Ovchinnikov, R. Barzilay, T. S. Jaakkola, F. DiMaio, M. Baek, D. Baker, De novo design of protein structure and function with RFdiffusion. *Nature*, 1–3 (2023).
44. J. Butcher, R. Krishna, R. Mitra, R. I. Brent, Y. Li, N. Corley, P. T. Kim, J. Funk, S. Mathis, S. Salike, A. Muraishi, H. Eisenach, T. R. Thompson, J. Chen, Y. Politanska, E. Sehgal, B. Coventry, O. Zhang, B. Qiang, K. Didi, M. Kazman, F. DiMaio, D. Baker, De novo Design of All-atom Biomolecular Interactions with RFdiffusion3. bioRxiv [Preprint] (2025). <https://doi.org/10.1101/2025.09.18.676967>.
45. H. Stark, F. Faltings, M. Choi, Y. Xie, E. Hur, T. O'Donnell, A. Bushuiev, T. Uçar, S. Passaro, W. Mao, M. Reveiz, R. Bushuiev, T. Pluskal, J. Sivic, K. Kreis, A. Vahdat, S. Ray, J. T. Goldstein, A. Savinov, J. A. Hambalek, A. Gupta, D. A. Taquiri-Diaz, Y. Zhang, A. K. Hatstat, A. Arada, N. H. Kim, E. Tackie-Yarboi, D. Boselli, L. Schnaider, C. C. Liu, G.-W. Li, D. Hnisz, D. M. Sabatini, W. F. DeGrado, J. Wohlwend, G. Corso, R. Barzilay, T. Jaakkola, BoltzGen: Toward Universal Binder Design. bioRxiv [Preprint] (2025). <https://doi.org/10.1101/2025.11.20.689494>.
46. J. Dauparas, I. Anishchenko, N. Bennett, H. Bai, R. J. Ragotte, L. F. Milles, B. I. M. Wicky, A. Courbet, R. J. de Haas, N. Bethel, P. J. Y. Leung, T. F. Huddy, S. Pellock, D. Tischer, F. Chan, B. Koepnick, H. Nguyen, A. Kang, B. Sankaran, A. K. Bera, N.

- P. King, D. Baker, Robust deep learning–based protein sequence design using ProteinMPNN. *Science* **0**, eadd2187 (2022).
47. L. E. McCoy, M. J. van Gils, G. Ozorowski, T. Messmer, B. Briney, J. E. Voss, D. W. Kulp, M. S. Macauley, D. Sok, M. Pauthner, S. Menis, C. A. Cottrell, J. L. Torres, J. Hsueh, W. R. Schief, I. A. Wilson, A. B. Ward, R. W. Sanders, D. R. Burton, Holes in the Glycan Shield of the Native HIV Envelope Are a Target of Trimer-Elicited Neutralizing Antibodies. *Cell Rep.* **16**, 2327–2338 (2016).
 48. A. J. Bahn-Suh, L. F. Caldera, P. N. P. Gnanapragasam, J. R. Keeffe, M. S. Seaman, P. J. Bjorkman, S. L. Mayo, Stabilized gp120-specific CD4 for next-generation HIV-1 inhibitors. bioRxiv [Preprint] (2026). <https://doi.org/10.64898/2026.03.24.713825>.
 49. J. S. McLellan, M. Chen, M. G. Joyce, M. Sastry, G. B. E. Stewart-Jones, Y. Yang, B. Zhang, L. Chen, S. Srivatsan, A. Zheng, T. Zhou, K. W. Graepel, A. Kumar, S. Moin, J. C. Boyington, G.-Y. Chuang, C. Soto, U. Baxa, A. Q. Bakker, H. Spits, T. Beaumont, Z. Zheng, N. Xia, S.-Y. Ko, J.-P. Todd, S. Rao, B. S. Graham, P. D. Kwong, Structure-Based Design of a Fusion Glycoprotein Vaccine for Respiratory Syncytial Virus. *Science* **342**, 592–598 (2013).
 50. D. Corti, J. Voss, S. J. Gamblin, G. Codoni, A. Macagno, D. Jarrossay, S. G. Vachieri, D. Pinna, A. Minola, F. Vanzetta, C. Silacci, B. M. Fernandez-Rodriguez, G. Agatic, S. Bianchi, I. Giacchetto-Sasselli, L. Calder, F. Sallusto, P. Collins, L. F. Haire, N. Temperton, J. P. M. Langedijk, J. J. Skehel, A. Lanzavecchia, A Neutralizing Antibody Selected from Plasma Cells That Binds to Group 1 and Group 2 Influenza A Hemagglutinins. *Science* **333**, 850–856 (2011).
 51. R. Volden, T. Palmer, A. Byrne, C. Cole, R. J. Schmitz, R. E. Green, C. Vollmers, Improving nanopore read accuracy with the R2C2 method enables the sequencing of highly multiplexed full-length single-cell cDNA. *Proc. Natl. Acad. Sci.* **115**, 9726–9731 (2018).

CONCLUSION AND OUTLOOK

5.1 Conclusion

The recurring challenge addressed in this thesis is that viruses responsible for global pandemics are often defined by their ability to escape narrowly focused immune responses. HIV-1 and SARS-CoV-2 differ in their biology, infection timelines, and evolutionary dynamics, yet both illustrate the same central problem. Viral surface proteins that mediate entry must preserve essential receptor-binding and fusion functions, but they do so while tolerating substantial sequence variation, glycan shielding, conformational masking, or rapid antigenic drift. As a result, effective countermeasures must identify and exploit regions of viral vulnerability that are conserved, functionally constrained, and accessible under the right structural or immunological conditions.

Chapter 2 addressed this problem in the context of sarbecovirus vaccine design. Building on prior mosaic nanoparticle studies, we used computational methods to select or engineer RBD antigens that reduced antibody recognition of variable class 1 and class 2 epitopes while preserving recognition of more conserved class 3 and class 4 epitopes (1–4). When displayed on mosaic nanoparticles, these antigens elicited cross-reactive antibody responses in mice. In particular, mosaic-7_{COM} outperformed earlier mosaic designs in naïve and pre-vaccinated mice, supporting the idea that computational antigen selection can improve the ability of mosaic nanoparticles to focus immune responses toward conserved sites of vulnerability.

Chapter 3 focused on a different but conceptually related strategy for HIV-1, using the virus's dependence on CD4 as an entry receptor against it. CD4-containing biologics are attractive because HIV-1 cannot easily escape CD4 binding without compromising viral fitness (5–7). However, prior CD4-Ig molecules have been limited by poor pharmacokinetics and off-target binding to class II MHC. By combining rational and computational design, we engineered gp120-specific CD4, or gCD4, which retained high-affinity gp120 binding while

reducing class II MHC binding and improving thermostability. When incorporated into CD4-Ig formats, gCD4 improved pharmacokinetic behavior and neutralization breadth, including activity against clinically relevant HIV-1 strains. These findings support gCD4 as a platform for next-generation HIV-1 entry inhibitors.

Chapter 4 then explored the discovery of nanobodies against conserved, receptor-induced surfaces on HIV-1 Env. The caldera represents a highly conserved and glycan-free region exposed upon CD4 engagement, but its recessed and transient nature makes it difficult to target with conventional antibodies. Using soluble Env trimers (SOSIPs), conformationally locked DS-SOSIPs, soluble CD4 (sCD4), 21c blocking conditions, yeast display, and iterative MACS and FACS enrichment, we established a screening framework for isolating open-state Env binders. Although this initial campaign did not identify caldera-specific nanobodies, it yielded candidates that preferentially bound open BG505-sCD4 relative to closed DS-BG505 and appeared to recognize a composite quaternary epitope involving both open Env and CD4. Thus, the screen succeeded in enriching for conformationally selective binders while revealing important constraints for future caldera-focused discovery efforts (8–13).

Taken together, these studies demonstrate three complementary molecular engineering approaches for countering viral diversity. Mosaic nanoparticles can bias humoral immunity toward conserved sarbecovirus epitopes. Engineered CD4 biologics can exploit HIV-1's conserved receptor-binding requirement while reducing liabilities associated with native CD4. Conformationally selective nanobody discovery can probe transient viral entry intermediates that are poorly sampled by natural antibody responses. Across all three projects, the central theme is that conserved viral vulnerabilities are often not simply exposed and immunodominant. They must be engineered into view, selectively stabilized, or discovered through screening strategies that account for viral structure, dynamics, and immune selection.

5.2 Outlook

The work described here points toward several future directions. For sarbecovirus vaccine design, mosaic-7_{COM} should be evaluated in additional animal models and against more recent SARS-CoV-2 variants and zoonotic sarbecoviruses. Future studies should also define whether the improved serum breadth observed in mice translates to durable protection, long-lived memory B cell responses, and protection against authentic virus challenge. At the design level, expanding the computational framework beyond available DMS datasets could enable the selection of antigens with larger numbers of mutations, improved developability, and better control over epitope accessibility. The broader principle of using antigenic diversity to disfavor strain-specific B cell activation may also be useful for other highly mutable viruses, including influenza and HIV-1.

For HIV-1 entry inhibition, gCD4 provides a promising platform for developing biologics that combine receptor mimicry with improved specificity and drug-like properties. Future work should evaluate gCD4-containing molecules in models that better capture human class II MHC expression and Fc-dependent effector functions. Additional formats could also be explored, including CD4-bNAb fusions and bispecific inhibitors that combine gp120 engagement with a second Env-directed specificity. Because gCD4 opens Env through the same receptor-binding pathway used during infection, it may also serve as a useful component in molecules designed to expose and then target otherwise hidden epitopes.

For anti-caldera binder discovery, future screening campaigns should reduce dependence on sCD4 and instead exclusively use the high-affinity CD4-mimetic M48U1 to open Env. This would simultaneously eliminate enrichment of anti-CD4 binders and provide more stable open SOSIPs during screens. Incorporating glycan-filled Env trimers, improved open-state probes, and more stringent negative selections should help distinguish binders to the desired protein surface from binders to strain-specific glycan holes or composite CD4-associated epitopes. In parallel, computational binder design methods may provide an alternative route to caldera binders that are difficult to recover from existing libraries. If such binders are identified, structural studies will be essential to confirm epitope specificity and determine

whether these molecules can neutralize HIV-1 directly or function as part of multivalent entry inhibitors.

More broadly, this thesis supports an approach to antiviral design in which molecular engineering is used to reshape the interaction between viral evolution and immune recognition. Rather than simply responding to each new variant or strain, future countermeasures can be designed around the constraints that viruses cannot easily escape. These constraints may take the form of conserved receptor-binding surfaces, conformational intermediates required for entry, or epitopes shared across diverse viral lineages. By combining structure-guided design, computational antigen selection, directed evolution, and conformationally selective screening, it should be possible to build vaccines and biologics that are more resilient to viral diversity and better prepared for future pandemics.

5.3 References

1. A. A. Cohen, P. N. P. Gnanapragasam, Y. E. Lee, P. R. Hoffman, S. Ou, L. M. Kakutani, J. R. Keeffe, H.-J. Wu, M. Howarth, A. P. West, C. O. Barnes, M. C. Nussenzweig, P. J. Bjorkman, Mosaic nanoparticles elicit cross-reactive immune responses to zoonotic coronaviruses in mice. *Science* **371**, 735–741 (2021).
2. A. A. Cohen, N. van Doremalen, A. J. Greaney, H. Andersen, A. Sharma, T. N. Starr, J. R. Keeffe, C. Fan, J. E. Schulz, P. N. P. Gnanapragasam, L. M. Kakutani, A. P. West, G. Saturday, Y. E. Lee, H. Gao, C. A. Jette, M. G. Lewis, T. K. Tan, A. R. Townsend, J. D. Bloom, V. J. Munster, P. J. Bjorkman, Mosaic RBD nanoparticles protect against challenge by diverse sarbecoviruses in animal models. *Science* **377**, eabq0839 (2022).
3. A. A. Cohen, J. R. Keeffe, A. Schiepers, S. E. Dross, A. J. Greaney, A. V. Rorick, H. Gao, P. N. P. Gnanapragasam, C. Fan, A. P. West, A. I. Ramsingh, J. H. Erasmus, J. D. Pata, H. Muramatsu, N. Pardi, P. J. C. Lin, S. Baxter, R. Cruz, M. Quintanar-Audelo, E. Robb, C. Serrano-Amatriain, L. Magneschi, I. G. Fotheringham, D. H. Fuller, G. D. Victora, P. J. Bjorkman, Mosaic sarbecovirus nanoparticles elicit cross-reactive responses in pre-vaccinated animals. *Cell* **187**, 5554–5571.e19 (2024).
4. E. Wang, A. A. Cohen, L. F. Caldera, J. R. Keeffe, A. V. Rorick, Y. M. Adia, P. N. P. Gnanapragasam, P. J. Bjorkman, A. K. Chakraborty, Designed mosaic nanoparticles enhance cross-reactive immune responses in mice. *Cell* **188**, 1036–1050.e11 (2025).
5. D. J. Capon, S. M. Chamow, J. Mordenti, S. A. Marsters, T. Gregory, H. Mitsuya, R. A. Byrn, C. Lucas, F. M. Wurm, J. E. Groopman, S. Broder, D. H. Smith, Designing CD4 immunoadhesins for AIDS therapy. *Nature* **337**, 525–531 (1989).
6. M. R. Gardner, L. M. Kattenhorn, H. R. Kondur, M. von Schaewen, T. Dorfman, J. J. Chiang, K. G. Haworth, J. M. Decker, M. D. Alpert, C. C. Bailey, E. S. Neale, C. H. Fellingner, V. R. Joshi, S. P. Fuchs, J. M. Martinez-Navio, B. D. Quinlan, A. Y. Yao, H. Mouquet, J. Gorman, B. Zhang, P. Poignard, M. C. Nussenzweig, D. R. Burton, P. D. Kwong, M. Piatak, J. D. Lifson, G. Gao, R. C. Desrosiers, D. T. Evans, B. H. Hahn, A. Ploss, P. M. Cannon, M. S. Seaman, M. Farzan, AAV-expressed eCD4-Ig provides durable protection from multiple SHIV challenges. *Nature* **519**, 87–91 (2015).
7. A. J. Bahn-Suh, L. F. Caldera, P. N. P. Gnanapragasam, J. R. Keeffe, M. S. Seaman, P. J. Bjorkman, S. L. Mayo, Stabilized gp120-specific CD4 for next-generation HIV-1 inhibitors. bioRxiv [Preprint] (2026). <https://doi.org/10.64898/2026.03.24.713825>.
8. J. P. Moore, J. A. McKeating, R. A. Weiss, Q. J. Sattentau, Dissociation of gp120 from HIV-1 Virions Induced by Soluble CD4. *Science* **250**, 1139–1142 (1990).

9. H. Wang, C. O. Barnes, Z. Yang, M. C. Nussenzweig, P. J. Bjorkman, Partially Open HIV-1 Envelope Structures Exhibit Conformational Changes Relevant for Coreceptor Binding and Fusion. *Cell Host Microbe* **24**, 579-592.e4 (2018).
10. K.-M. A. Dam, C. Fan, Z. Yang, P. J. Bjorkman, Intermediate conformations of CD4-bound HIV-1 Env heterotrimers. *Nature* **623**, 1017–1025 (2023).
11. P. Acharya, T. S. Luongo, M. K. Louder, K. McKee, Y. Yang, Y. Do Kwon, J. R. Mascola, P. Kessler, L. Martin, P. D. Kwong, Structural Basis for Highly Effective HIV-1 Neutralization by CD4-Mimetic Miniproteins Revealed by 1.5 Å Cocrystal Structure of gp120 and M48U1. *Structure* **21**, 1018–1029 (2013).
12. C. McMahon, A. S. Baier, R. Pascolutti, M. Wegrecki, S. Zheng, J. X. Ong, S. C. Erlandson, D. Hilger, S. G. F. Rasmussen, A. M. Ring, A. Manglik, A. C. Kruse, Yeast surface display platform for rapid discovery of conformationally selective nanobodies. *Nat. Struct. Mol. Biol.* **25**, 289–296 (2018).
13. S. Muyldermans, Nanobodies: Natural Single-Domain Antibodies. *Annu. Rev. Biochem.* **82**, 775–797 (2013).