Complete mapping of HIV-1 escape from broadly neutralizing antibodies, vaccines, and drugs

Adam S. Dingens

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2019

Reading Committee:
Jesse Bloom, Chair
Julie Overbaugh
Kelly Lee

Program Authorized to Offer Degree:
Molecular and Cellular Biology

University of Washington

Abstract

Complete mapping of HIV-1 escape from broadly neutralizing antibodies, vaccines, and drugs

Adam S. Dingens

Chair of the Supervisory Committee:
Affiliate Professor Jesse Bloom
Associate Member, Division of Human Biology, Fred Hutchinson Cancer Research Center

The expansive global diversity of HIV-1 Env presents significant hurdles in developing a broadly protective vaccine. This diversity is a result of HIV Env's exceptional evolutionary capacity, which allows it to evade the extraordinary diversity of the humoral immune system during infection. However, the evolutionary arms race between Env and humoral immunity occasionally drives the development of broadly neutralizing antibodies (bnAbs) capable of neutralizing diverse strains. Mapping the epitope specificity of bnAbs has revealed conserved regions of Env, which are promising targets for structure-based vaccine design. Additionally, bnAbs' broad activity and potential to direct the killing of infected cells make them promising antiviral immunotherapeutic drugs for HIV prevention, therapy, and cure strategies. Translating bnAbs into vaccines and therapies will require both a detailed understanding of how bnAbs interact with Env as well as assessing their potential for viral escape.

While structural studies provide atomic-level views of HIV-antibody interactions, they fail to reveal the functional interactions necessary for neutralization and the viral mutations that disrupt these interactions. Neutralization and binding assays using mutants can provide such information for specific mutations, but even the largest studies employing one-at-a-time mutagenesis can only assay a small fraction of all possible Env mutations.

To overcome these shortcomings, we have developed *mutational antigenic profiling*, a deep mutational scanning approach that completely maps the functional interface between HIV and an antibody in a single massively parallel experiment. This involves generating libraries of HIV that carry all possible amino-acid mutations to Env (12,730 amino-acid mutations), incubating these viral libraries with or without an antibody, infecting T cells, and using deep sequencing to quantify the enrichment of each mutation in the antibody selected versus non-selected libraries. Profiling escape from bnAb PGT151 identified all previously known and revealed numerous additional escape mutations. Benchmarking these data against traditional neutralization assays further validated that we accurately quantified the effect of all amino-acid mutations to Env. Additionally, evaluating the effect of each amino acid at each site elucidated the biochemical mechanisms of escape throughout the epitope, highlighting the previously unappreciated role for charge-charge repulsions.

To gain a broad view of HIV antibody escape, we mapped escape from a panel of nine bnAbs targeting the five best-characterized Env epitopes. Importantly, many of these bnAbs are being clinically developed as immunotherapeutics. While prior studies had defined each of these bnAbs' *structural epitope,* our unbiased mapping defined their *functional epitopes,* or the sites at which mutations mediated escape in the context of replication competent viruses, for the first time. For most bnAbs, mutations at only a small fraction of structurally defined contact sites

mediated escape, and escape often occurred at sites that are near but do not directly contact the antibody. Further, these data helped to interpret viral mutations observed in immunotherapy clinical trials—*in vivo* escape occurred in the functional epitope, some of which was previously missed since it was far from the structural epitope. Additionally, this data allowed for an unbiased quantification of the ease of viral escape for each bnAb, which we found is distinct from antibody breadth.

We also mapped escape from a pool of two bnAbs; we found that there were no mutations that robustly escaped both antibodies, agreeing with the results of two recently completed clinical trials that administered this combination. Further, we profiled escape from two antibodies across multiple viral strains, providing the first unbiased quantifications of strain-specific differences in antibody escape.

Next, we leveraged mutational antigenic profiling to directly refine structure-based vaccine design. We contrasted escape from bnAb VRC34.01 with escape from two murine antibodies that were elicited with immunogens based on the VRC34.01 epitope. This revealed distinct differences in the recognition of natural and vaccine-elicited antibodies, and provide a template to guide the iterative rounds of vaccine design.

We then adapted this approach to better delineate the genotypic determinants of resistance to the only clinically approved HIV fusion inhibitor, enfuvirtide. Again, we identified both previously characterized and novel resistance mutations. Many resistance mutations were allosteric to the drug's binding site, which shed light on diverse mechanisms of resistance. Further, this complete map of resistance may be of use in the clinical monitoring of resistance during therapy and the genotypic prediction of enfuvirtide sensitivity prior to treatment.

Few protein-protein interfaces have been as heavily studied as those between bnAbs and Env, as these interactions provide the motivation for many HIV treatment and prevention efforts. Mutational antigenic profiling yields an unprecedented view of these interfaces, redefining out understanding of an antibody's functional epitope. The complete maps of viral escape detailed in this thesis provide a mutation-level antigenic atlas for understanding viral immune escape and guiding the development of antibody immunotherapies and vaccines.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ADCC: antibody-dependent cell-mediated cytotoxicity

AIDS: acquired immunodeficiency syndrome

ART: antiretroviral therapy

bnAb: broadly neutralizing antibody

CD4bs: CD4 binding site

CDR: complementarity determining region

CHR: C-terminal heptad repeat region

ELISA: enzyme-linked immunosorbent assay

Env: envelope

HCDR3: heavy chain 3 complementarity determining region

HIV: human immunodeficiency virus

HR2: heptad repeat 2 region

IC_{50}: inhibitory concentration at which 50% of the input virus is neutralized

IgG: immunoglobulin G

mAb: monoclonal antibody

MPER: membrane proximal external region of HIV Env

nAb: neutralizing antibody

NHR: N-terminal heptad repeat region

PNG: potential N-linked glycosylation site

SHIV: simian/human immunodeficiency virus

SHM: somatic hypermutation

V1/V2: variable regions 1 and 2 of HIV Env

V3: variable region 3 of HIV Env

VH: IgG heavy chain variable domain

VL: IgG light chain variable domain

ACKNOWLEDGEMENTS

Thanks to Jesse and Julie, for giving me every possible opportunity to succeed and more. Thanks for the freedom to chase my own ideas and interests, and for reeling me back in when these ideas inevitably wandered too far. Thanks for caring about much more than just my science, and for being incredibly supportive and understanding when contemplating my next career steps. Thanks to Julie for investing in me as both a scientist and a mentor, and specifically for giving me the opportunity to work with Haidyn and Dana. Thanks to Jesse for always giving me more ownership in projects than I deserve, and for pushing me to help guide some of the lab's longer-term goals. The most valuable training I have received is simply watching Julie and Jesse take different approaches to leading diverse people and projects. I am incredibly grateful to have these two role models to look up to and emulate.

The amazing people Jesse and Julie have attracted to their labs also speaks to their strengths. This large group has obviously made me a better scientist, but their friendship and personal advice through the highs and lows of both grad school and life has sincerely made me a better person. Thanks for the banter, inspiration, practice talks, paper edits, lab meetings with far too many tangents and questions (I apologize), and the donuts. And thanks for important things that got me through grad school - the hiking, climbing (two volcano summits with labmates!), backpacking, and skiing. A sincere thank you to each and every one of you.

Thanks to Hugh Haddox, who first developed the HIV Env deep mutational scanning techniques that I leveraged in this thesis. But even more important was his patient mentorship, support, and friendship. Working alongside Hugh in the HIV Corner and continually bouncing ideas back and forth is one of the things in my PhD that I look back at most fondly. Now when it is my turn to mentor labmates, I try to channel my inner Hugh – if I can provide even a fraction of the personal and intellectual inspiration that Hugh did to me, I will be incredibly proud.

Thank you to Mike Doud, who developed the analytical frameworks for the bulk of the data analyses in this thesis. Mike's work and advice certainly helped shape much of my own work. Juhye Lee seamlessly played a similar role in pushing forward both concepts and technologies in mapping escape from antibodies. It is easy to do good science when following Juhye and Mike's lead!

Thanks to Sarah Hilton, an honorary (full?) member of the HIV Corner, for her emotional and computational support. Thanks to Jeremy Roop, Kate Dusenbury, and Caelan Radford for graciously listening to my long-winded opinions; working with and learning from each of you has been a pleasure. Thank you to Dana Arenz and Haidyn Weight, for assistance in validating much of the data this thesis. Also thanks to Dana and Haidyn for their patience and independence, and for helping me become a better mentor.

Thank you to Keshet Ronen for welcoming me into a project she carefully developed and for providing excellent analytical training. Thanks to Stephanie Rainwater for help on Keshet and I's work, and for the careful and patient HIV room training. Thank you to Cassie Simonich and Laura Doepker, as well as a larger network of collaborators, including Duncan Ralph, Eric Matsen, James Williams, and Kelly Lee, for rapidly sharing reagents, data, and advice. Having the opportunity to play off of this exceptional team's work has made it incredibly easy to pursue exciting ideas.

While not covered in this thesis, I have had the precious opportunity to work with some invaluable patient samples from Kenyan cohorts, enabled by Julie's incredible foresight and commitment to long-term international collaborations. Thank you to the women and infants who enabled this work by donating their time and samples. Also thanks to the epidemiological collaborators in Seattle and Kenya, as well as the clinic and lab staff who keep these cohorts running smoothly.

Thank you to the many other faculty at the Hutch and UW who contributed to my training. Specifically, thanks to Susan Graham for mentoring my Epidemiology MS project and for bestowing a very different set of epidemiology and writing skills than those I learned from other advisors. Thanks to Dara Lehman, for being a third advisor when two opinions were not enough. Thanks to Michael Emerman for providing critical insights into experimental details, for giving frank career advice, and for teaching my favorite class at any level. Also thanks to Michael for organizing the Thursday Morning Virology Group meetings, and to all of it's participants. The feedback, diverse topics, and open, collaborative atmosphere of these meetings have shaped my work and scientific viewpoint. Thanks also to members of my thesis committee – Kelly Lee, Doug Fowler, and Shiu-Lok Hu, and particularly Kelly for also being on my reading committee.

Thank you to the many collaborators who strengthened and extended my work, including those from the Trkola, Kwong, Mascola, and Gilbert groups. A special thanks to Alexandra Trkola, Matthias Glögl, and Nikolas Friedrich for hosting me in beautiful Zürich.

Thank you to Alex Moreno, Shama Samant, and the MCB program for excellent administrative support. Also thanks to Shama for being my adventurous travel partner on a safari in Tsavo!

Thank you to Bruce and Jacqueline Cook, and the Cook Family Foundation. I almost certainly would not be where I am today without their generous support of my undergraduate education.

Thank you to the close friends from grad school who celebrated the good, commiserated the bad, and contemplated the future. Justin, Laura, Tiffany, Harley, and many labmates – I am very glad grad school brought us together.

Thanks to my parents, Ann and Jed, for getting me here. Thank you for instilling values of independence, perseverance, and wonder that led to this career path, and for making sure I have always had every opportunity to grow and learn. Thank you for your unwavering support, which has kept me going through graduate school. Thanks you also to my siblings – John Paul, Brandon, Ana, Ava, Chloe, and Eli. Each of you inspire me in your own way. Thanks for visiting me from across the country, and for being my biggest cheerleaders during grad school.

Lastly, thank you to Jacky. I am incredibly thankful for your support and encouragement throughout this long journey, which you've made far more enjoyable. Thank you for both lifting me up and keeping me grounded on a near daily basis. And of course, thank you for being my adventure partner throughout the Cascades and across the world.

Chapter I

Introduction

The HIV-1/AIDS pandemic

HIV-1/AIDS continues to be one of the greatest global health burdens thirty-five years after its discovery, with an estimated 36.9 million people infected and 940,000 deaths annually [1]. Despite significant advances in treatment and prevention, including promising advances in using antiretrovirals as pre-exposure prophylaxis and treatment as prevention [2], there are approximately 1.8 million new infections annually [1], concentrated in some of the world's most disadvantaged and marginalized communities.

These sobering statistics highlight the dire need for a preventative vaccine. While few topics have received as much research as an HIV-1 vaccine, the virus has been remarkably recalcitrant to these efforts. The virus's extensive antigenic diversity, highlighted by the nine distinct viral subtypes that can share as little as 64% identity in the immunodominant viral Env protein [3], presents one of the starkest difficulties in designing a broadly protective antibody-based vaccine.

HIV-1 Envelope

Env is the only viral protein present on the surface of virions and is the target for neutralizing antibody responses. This trimeric receptor glycoprotein, made up of three surface (gp120) and transmembrane (gp41) heterodimers, is responsible for entry into host cells. Env first binds to the cellular receptor CD4, which results in conformationally rearrangement and

"opening" of the trimer, creating the CCR5 or CXCR co-receptor binding site [4]. Subsequent co-receptor binding triggers conformational rearrangements in gp41, resulting the insertion of the fusion peptide into the target cell membrane and the formation of the putative "pre-hairpin intermediate" structure [5]. Finally, collapse of gp41's two heptad repeat regions forms the post-fusion six helical bundle, providing the energy necessary to fuse the target cell and virion membranes, facilitating viral entry [6].

Env's expansive genetic and antigenic diversity is the result of HIV's exceptional ability to evade immunity through rapid sequence evolution, facilitated by its large population size, quick generation time, and error-prone replication. Further, Env has evolved many additional immune evasion strategies. Env is covered by a host-derived "glycan shield" that encompasses half of its mass and cloaks a majority of the protein surface [7], and virions evades antibody cross-linking avidity effects by limiting trimer density on the virus surface [8]. Additionally, Env is conformationally dynamic and only exposes certain epitopes transiently [9] or post receptor binding [10], and the virus displays decoy epitopes that elicit non-neutralizing antibodies via presentation of nonfunctional Env conformations and "stumps" [11].

HIV-1 antibodies

After infection, antibodies to immunodominant variable loops develop within the first few weeks [12]. However, early binding antibody responses fail to neutralize the virus and exert little selective pressure on Env [13]. After a few months, neutralizing antibodies develop, but are often specific to the autologous virus, with limited heterologous breadth, meaning they fail to neutralize other diverse viral strains [12]. While these antibodies do exert selective pressure on Env, the viral population is able to evolve resistance [14–16], staying one step ahead of the

developing antibody response *in vivo*. While some of my previous work illustrated how antibody responses found during infection likely do not play a role in protecting against superinfection (reinfection with a distinct viral strain) [17], characterizing exceptional antibody responses may still provide a roadmap to HIV prevention.

Over the last decade, a burgeoning number of broadly neutralizing antibodies (bnAbs) have been isolated from HIV-infected humans. The isolation and characterization of these bnAbs, found in 10-50% of individuals depending on the definition of "broad", has identified functionally and sequence conserved Env epitopes that are promising vaccine targets [18]. These sites include the CD4 binding site (CD4bs), membrane proximal external region (MPER), N332 glycan supersite, V1/V2 trimer apex, and gp120-gp41 interface/fusion peptide [18].

BnAbs, which usually arise after multiple years of infection and iterative rounds of viral escape and subsequent affinity maturation [19], often have unique genetic features. In addition to their extensive somatic hypermutation (up to ~30% form germline due to affinity maturation), bnAbs often have long HCDR3 loops and particular antibody variable gene usage [18]. The unique sequence features of bnAbs allow them to penetrate the glycan shield and target conserved Env epitopes. For example, the long, rigid HCDR3 loop of PGT145 allows it to reach through the glycans at the trimer apex and bind an epitope at the trimer interface [20]. Further, the breadth of CD4bs bnAbs correlates with their angle of approach [21], which likely reflects their ability to access the CD4bs while accommodating steric hindrance from neighboring glycans [7]. This unique targeting of the CD4bs is driven by either the usage of a particular VH gene or HCDR3 rearrangement [21], highlighting how bnAbs can find general approaches to thwart Env's immune evasion tactics and access conserved epitopes.

BnAb-based vaccines and immunotherapies

A detailed characterization of bnAb epitopes has been the basis for structure- or epitope-based vaccine design. Eliciting responses to such sites has thus far had limited success [22], likely due to the highly specific ontogeny, binding, and circumvention of Env's defenses that each bnAb employs. However, structure-based vaccine design is starting to show its true potential for developing a broadly protective HIV vaccine [23–26]. Admittedly, any vaccination approach that tries to mimic the prolonged antigenic exposure that results in the development of bnAbs has many hurdles to overcome, but there have recently been promising advances in bnAb priming and sequential immunization strategies that aim to provoke the intricate affinity maturation processes of bnAbs [24,26–28]. In order to elicit such highly specific responses with vaccination, a more detailed understanding of bnAb epitopes and the determinants of bnAb elicitation is needed. This includes (but is certainly not limited to) high-resolution definitions of exactly what residues encompass an epitope, paraptope-epitope binding chemistries, mechanisms of neutralization, and identification of allosteric sites that may influence epitope exposure via altering steric hindrance or conformational dynamics.

In addition to revitalizing antibody-based vaccine efforts, the identification of bnAbs has also allowed for the development of novel antibody-based therapies. bnAbs' broad activity and potential to direct the killing of infected cells through Fc-mediated interactions make them ideal antiviral drugs for HIV prevention, therapy, and cure strategies.

Proof-of-concept studies in animal models have shown that passively transferred bnAbs can prevent infection [29–35], and suppress viremia during infection [36–40]. Further, administration of bnAbs to macaques has been shown to halt SHIV infection in infants [41] and delay or eliminate viral rebound after ARV discontinuation [42]**.** bnAb immunotherapies are also

rapidly progressing in humans, where they have resulted in the transient reduction of viral load or delay of viral rebound after treatment interruption [43–47]. Currently, there are clinical trials planned or underway to evaluate bnAbs as part of "shock and kill" strategies, to reduce reservoir size and/or allow "drug holidays" via administration at treatment interruption, to therapeutically suppress viremia, to prophylactically prevent infection (largely as a proof-of-concept), and to prevent infection in infants [48]. This rapidly advancing field of human clinical trials is detailed in Figure 1.1.

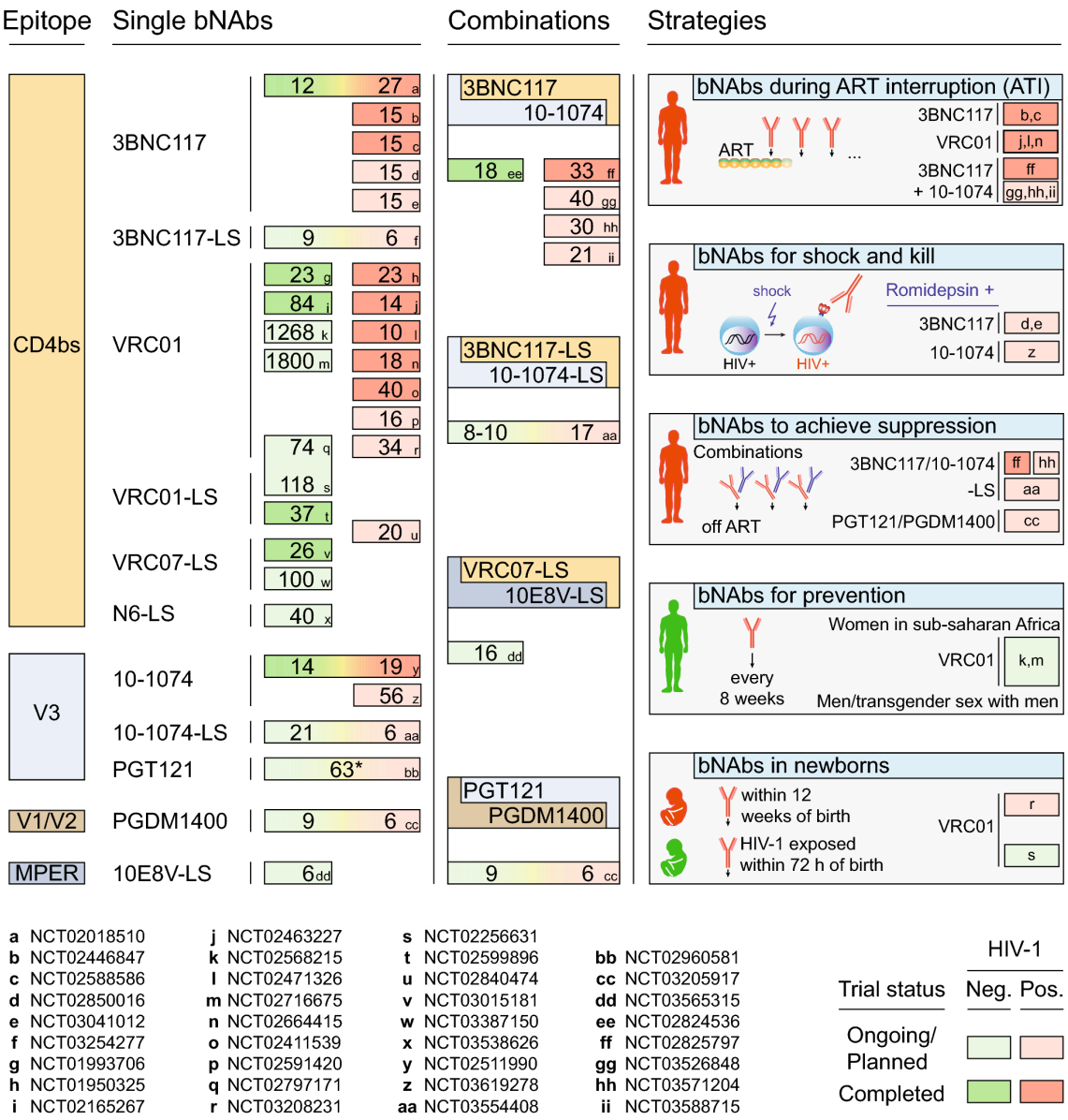


Figure 1.1. The clinical trial landscape of HIV bnAb interventions. Numbers indicate the number of trial participants receiving bnAbs, letters annotate the NIH Clinical Trial Registry Number for each trial. Adapted from [48].

However, bnAbs have an exceptionally difficult target to wrangle. HIV Env's exceptional evolutionary capacity allows the virus to stay one step ahead of bnAbs during infection, and resistance often arises when bnAbs are therapeutically administered to infected animals

[38,39,49] or humans [43–47,50]. Defining the viral mutations that mediate viral escape is therefore essential to optimizing and evaluating bnAb-based immunotherapies and vaccines. This is not always an easy task, as many mutations arise during infection and it is oftentimes not immediately clear which ones play a role in immune evasion. This then must be validated by testing individual mutations one-at-a-time in neutralization or binding experiments.

While massive efforts have been directed at mapping the epitope of bnAbs using structural and functional approaches, these data fail to completely define the *functional epitope*, defined as the sites where viral mutations can mediate escape from antibody neutralization. While bnAb escape mutants have been identified through a variety of experimental and observational means, we have yet to define all possible pathways of escape from even a single bnAb. Additionally, identification of all the amino-acids that mediate escape at a given residue, rather than just single mutants that have arisen during *in vivo* or *in vitro* antibody selection, would give clues as to the biochemical properties that mediate escape. This comprehensive view of escape is not easily achievable with current antibody mapping approaches, which are hindered by the complexities of Env-antibody interactions.

High-resolution structural studies, such as cryo-EM and crystallography, are often considered the gold standard for epitope mapping, but are labor-intensive and technically challenging to perform in the context of conformationally relevant Env. Further, even atomic-level structural characterization of Env-antibody interactions does not define the functional interactions that drive neutralization. Fine-resolution functional mapping currently relies on interrogation of single selected mutant pseudoviruses in neutralization assays, which is also labor-intensive, inherently biased, and does not reveal if escape mutants are actually viable. Finally, escape mutants have been identified through observational studies of bnAb escape in

infected individuals, animal models, or *in vitro* passaging, but these approaches only identify a small subset of all possible escape mutants due to limitations in codon sequence space, stochasticity, and selective sweeps.

Deep mutational scanning of HIV-1

To address shortcomings in existing approaches to study antibody epitopes and viral immune escape, we have developed deep mutational scanning approaches to interrogate the functional and antigenic effect of all possible amino acid mutations to Env in the context of replication competent viruses. Prior work in the Bloom lab developed this system to study the inherent mutational tolerance of HIV Env during viral replication in cell culture, which revealed that the epitope of bnAbs are inherently less tolerant to mutations, rigorously validating a pervasive idea in the field [51]. We further quantified the mutational tolerance in two additional transmitted/founder, subtype A Envs, and characterized how a the mutational tolerance of a small subset of sites has substantially shifted between these two Env homologs [52].

During my dissertation work, I have extended these deep mutational scanning approaches to quantify the effect of all functionally tolerated mutations to Env on antibody neutralization, in a process we term *mutational antigenic profiling*. In the following chapters, I will describe a proof-of-concept study detailing the validation and advantages of mutational antigenic profiling (**Chapter II)** [53], use it to evaluate a panel of bnAbs being developed as immunotherapeutics (**Chapter III)** [54], refine structure-based vaccine design (**Chapter IV**) [55], and completely profile resistance to an clinically-used antiretroviral drug (**Chapter V**) [56].

Chapter II

Comprehensive Mapping of HIV-1 Escape from a Broadly Neutralizing Antibody

The text in this chapter has been modified slightly from: Cell Host & Microbe 2017 Jun 14; 21 (6), 777-787. e4. doi: 10.1016/j.chom.2017.05.003

Summary

Precisely defining how viral mutations affect HIV's sensitivity to antibodies is vital to the development and evaluation of vaccines and antibody immunotherapeutics. But despite great effort, a full map of escape mutants has not yet been delineated for even a single anti-HIV antibody. Here we describe a massively parallel experimental approach to quantify how all single amino-acid mutations to Envelope (Env) affect HIV's sensitivity to a neutralizing antibody. We applied this approach to PGT151 and identified novel sites of escape in addition to those previously defined by structural and functional studies, such as glycans at sites 611 and 637, residue 647, and sites in the fusion peptide. Evaluating the effect of each amino acid at each site lends insight into the biochemical basis of escape throughout the epitope. Thus, comprehensive mapping of HIV antibody escape gives a quantitative, mutation-level view of the ways that Env can evade neutralization.

Introduction

HIV's rapid evolution enables it to outpace even the exceptional adaptive capacity of the humoral immune system. Although the immune system is usually outmatched in this evolutionary arms race, leading to viral persistence, infected individuals occasionally develop antibodies capable of neutralizing diverse viral strains. While these broadly neutralizing antibodies (bnAbs) do not control infection in the individual in whom they arise, their identification has motivated efforts in rational vaccine design and antibody immunotherapeutics. For example, epitope mapping of bnAbs has revealed conserved sites of vulnerability on HIV's envelope glycoprotein (Env), and a leading vaccine strategy is to design immunogens that elicit an antibody response targeting these conserved sites [18]. bnAbs are also being tested in both prophylactic and therapeutic settings. Numerous studies in animal models have shown proof-of-concept that passively infused bnAbs can protect against infection [29–35] and therapeutically suppress viremia during infection [36–40]. Similar bnAb-based immunotherapies are being tested in humans, with some early studies showing a transient reduction in viral load or delay of viral rebound after treatment interruption in some individuals [43–47].

Despite the impressive breadth and potency of bnAbs *in vitro,* HIV can eventually evade them *in vivo*. Viral isolates from individuals who develop bnAbs are typically resistant to neutralization, and resistance arises when bnAbs are administered to infected animal models [38,39,49] or humans [43–47,50]. It is therefore important to prospectively identify all *env* mutations that affect the sensitivity to a bnAb. However, this can be challenging, in part because bnAbs often target complex conformational and glycosylated epitopes. To date, a complete set of HIV escape mutations has yet to be elucidated for any antibody.

The limited observational studies of viral escape from bnAbs to-date likely reveal only a fraction of the full repertoire of escape mutations. Structural studies provide atomic-level views of the antibody-antigen footprint, but fail to reveal which interactions are necessary for neutralization and which mutations disrupt these interactions. Indeed, it has long been appreciated that binding energetics are often concentrated at select sites in the protein-protein interface [57,58], and mutations at Env residues that participate in crystal-structure-defined interactions do not always affect bnAb binding and neutralizing [59,60]. Because structures do not functionally define escape mutations, researchers often generate and interrogate single amino-acid mutants in binding or neutralization assays. This approach is so labor intensive that it has only been applied to a fraction of the sites in Env, and typically to only one or a few mutations – often to alanine – at these sites.

We have applied a new deep mutational scanning-based approach to comprehensively map all mutations to Env that enable HIV to escape from a broadly neutralizing antibody. This approach, mutational antigenic profiling, involves creating libraries of all single amino-acid mutants of Env in the context of replication-competent HIV [51], selecting for mutations that promote antibody escape, and using deep sequencing to quantify the enrichment of each mutation. It is conceptually similar to a strategy recently used by Doud *et al.* [61] to completely map the escape of influenza A from anti-hemagglutinin antibodies. Our work comprehensively profiles HIV escape from bnAb neutralization in the context of actual virus, distinguishing it from previous deep mutational scanning experiments that have used yeast or cell-surface display to identify Env variants that bind germline-encoded bnAb progenitors [23,62]. The map of escape mutations that we generate provides a complete view of the functional interface between

Env and a bnAb, and suggests biochemical mechanisms of escape that are not readily apparent from the incomplete data provided by traditional mapping approaches.

Results

Overview of mutational antigenic profiling

To define all single amino-acid mutations to Env that increase resistance to antibody neutralization, we implemented the mutational antigenic profiling approach in Figure 1. This approach relies on creating libraries of full-length proviral clones encoding all possible mutations in *env*, and passaging these libraries in T cells [51]. The result is a library of replication-competent HIV encoding all functionally tolerated mutations in Env. This mutant library is then incubated with and without antibody and used to infect target cells. After entry, the frequency of each mutation in the infected cells is quantified by deep sequencing (Figure 2.1A). The *differential selection* that the antibody exerts on a mutation is defined as the logarithm of that mutation's enrichment in the antibody-selected condition relative to the control, and selection for enriched mutations is plotted in logoplots as shown in Figure 2.1B. The entire experiment was performed in biological triplicate beginning from independent generation of the proviral plasmid mutant libraries (Figure 2.1C).

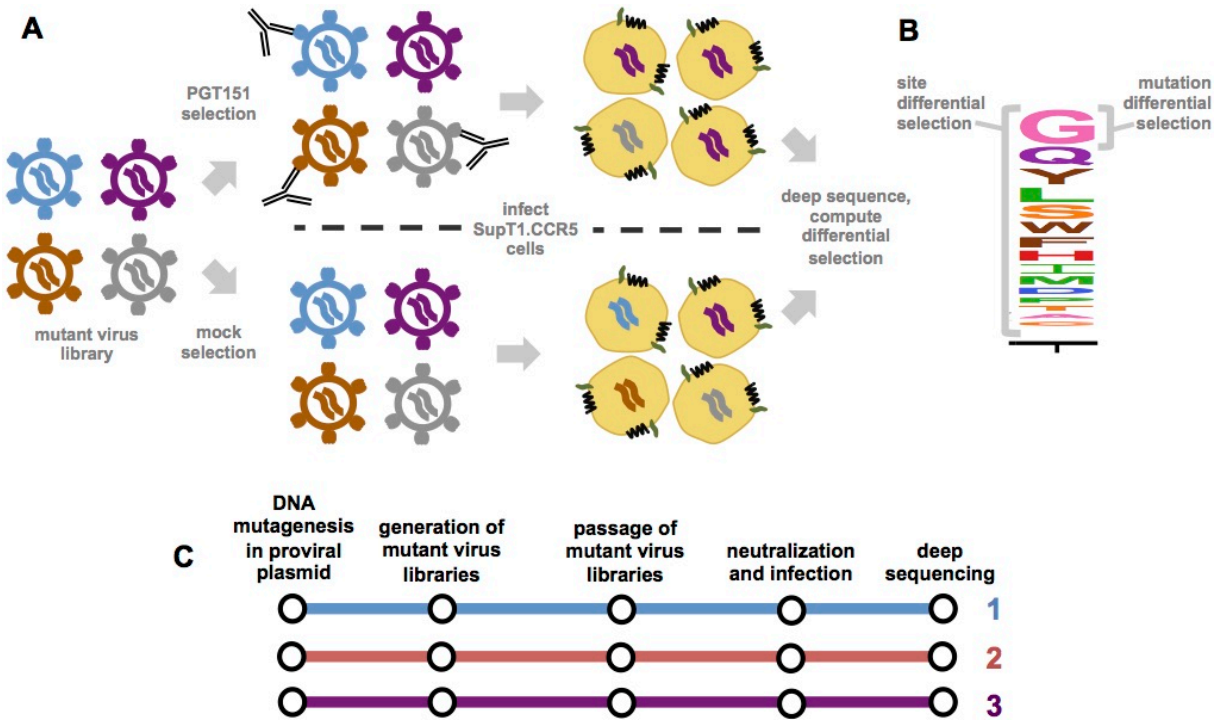


Figure 2.1. Schematic of mutational antigenic profiling.

A. Mutant virus libraries of HIV, which have been passaged in SupT1.CCR5 cells and thus should encode all single amino-acid mutants to Env compatible with viral replication, are incubated with and without an anti-HIV antibody, and then infected into SupT1.CCR5 cells. After viral entry and reverse transcription, cDNA is isolated and *env* is deep sequenced. The differential selection exerted by the antibody is quantified as the logarithm of the frequency of each mutation relative to wildtype in the antibody-selected condition compared to the control condition. **B**. A logoplot visualizing the differential selection at a single site. The height of each letter is proportional to the differential selection for that amino acid mutation. The site differential selection is the sum of all mutation differential selection values at that site. **C**. The entire mutational antigenic profiling process was completed in biological triplicate, starting from generation of three independent Env mutant libraries in the context of proviral plasmids.

Generation and deep sequencing of mutant virus libraries

We applied mutational antigenic profiling to a virus relevant to HIV transmission by generating the mutant libraries in the context of an *env* (BF520.W14M.C2) from a subtype A virus isolated shortly after mother-to-child transmission in an infant who went on to rapidly develop a bnAb response [63–65]. We introduced all possible codon-level mutations to the gp120 and extracellular portion of gp41 (Env residues 32 to 703) in the context of a proviral plasmid. We did not mutagenize Env's signal peptide or cytoplasmic tail, as these are not direct targets of neutralizing antibodies but play a role in regulating Env cell surface expression [66–68]. Sanger sequencing of individual proviral clones revealed that the number of codon mutations per clone followed a Poisson distribution with a mean of 1.1 mutations/clone, and the mutations were relatively uniformly distributed along the length of Env (Figure 2.2 and Figure 2.3).

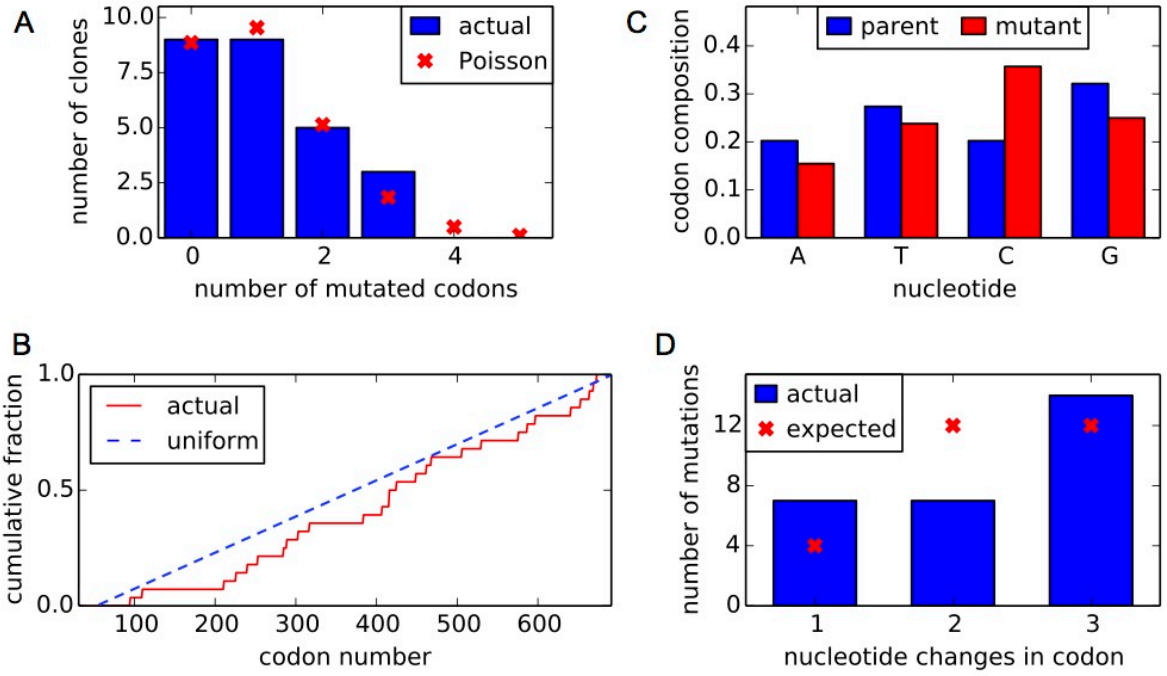
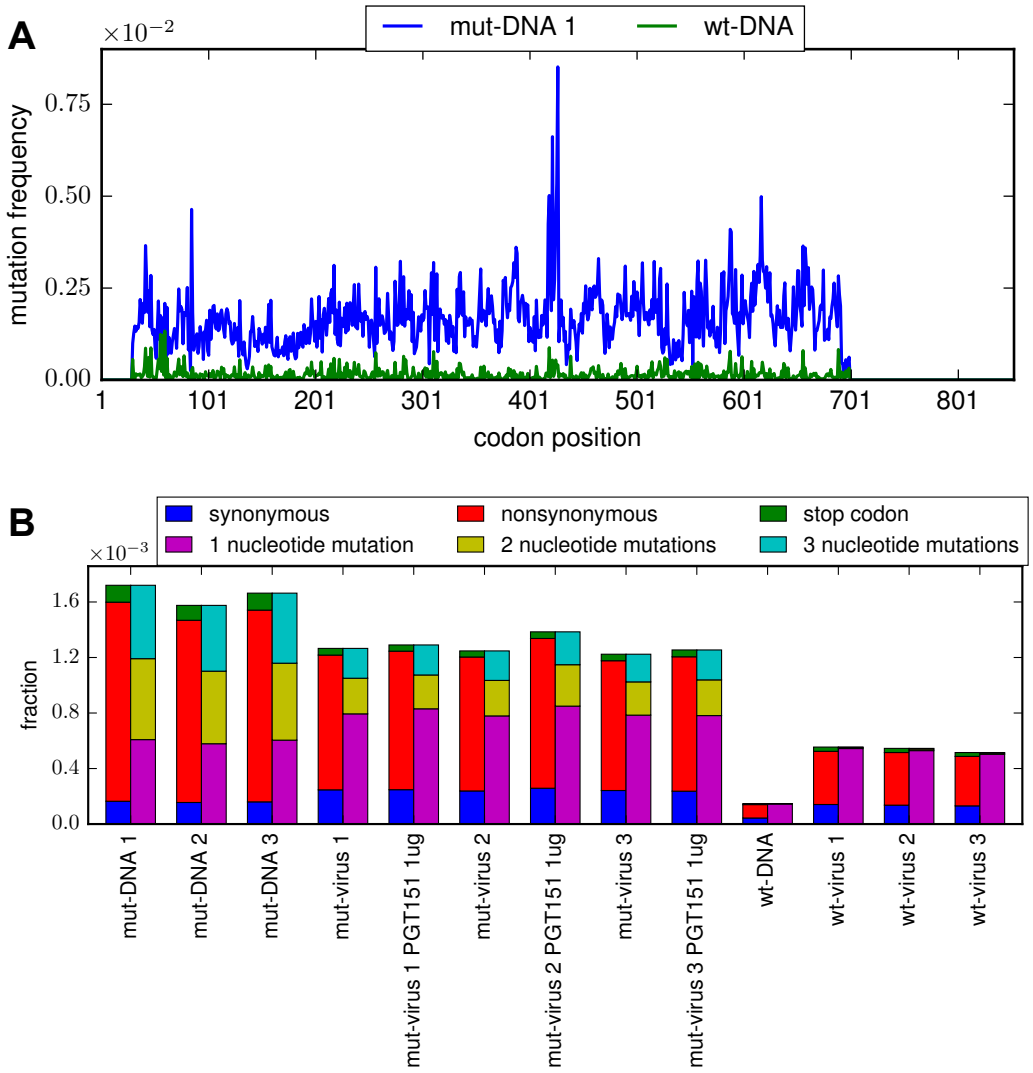


Figure 2.2. Mutant DNA library statistics based on Sanger sequencing data from 26

individual clones. Clones were obtained from the three replicate mutant libraries of the proviral plasmid. **A** The distribution of codon mutations per clone is shown, comparing the actual number with the expect number based on a Poisson distribution. **B.** The cumulative distribution of mutations across the length of the mutagenized portion of the gene compared to the expected uniform distribution. In this panel, codons are numbered sequentially along the Env sequence beginning with 1 at the N-terminal methionine. **C.** The nucleotide composition of parent and mutant codons for each mutation. The introduced mutations had relatively uniform nucleotide composition. **D.** The average number of nucleotide changes of each codon mutation compared to the expected if each codon is equally likely to be mutated to every other codon. Our codon mutagenesis introduced a mix of 1-nucleotide (e.g., GCA -> GtA), 2-nucleotide (e.g., GCA -> Gtc), and 3-nucleotide (e.g., GCA -> atc) codon mutations.



C

Sample	Total reads	Total aligned barcodes	Median effective depth
mut-DNA 1	13467833	1936861	234797
mut-DNA 2	15532281	2218771	310252
mut-DNA 3	12458983	1841391	234028
mut-virus 1	12489665	1127841	166493
mut-virus 1 PGT151 1ug	9549563	733525	85273
mut-virus 2	14157253	1671121	248043
mut-virus 2 PGT151 1ug	7135403	1069792	137706
mut-virus 3	10756401	1585321	188920
mut-virus 3 PGT151 1ug	4839760	739364	97915
wt-DNA	13290115	2011619	228442
wt-virus 1	9314434	1265066	183137
wt-virus 2	13885113	1764646	236349
wt-virus 3	11032360	1627549	227152

Figure 2.3. Mutation frequencies and sequencing depth of mutational antigenic profiling samples. A. The codon mutation frequency across the gene is, on average, >10 times higher in

the mutant DNA libraries (*mut-DNA*) than the error frequency observed when sequencing wildtype proviral plasmid (*wt-DNA*). In this panel, codons are numbered sequentially along the Env sequence beginning with 1 at the N-terminal methionine. **B.** The frequency of mutation types in the wildtype and mutant DNA and virus libraries. Compared to wildtype proviral plasmid (*wt-DNA*), there is an increase in single nucleotide mutations in the wildtype passaged virus libraries (*wt-virus*), presumably from errors during viral replication. There is also evidence of purging of deleterious mutations (e.g. stop codons) from the mutant plasmid libraries (*mut-DNA*) relative to the mutant viruses (*mut-virus*) grown from these plasmids, both in the presence and absence of PGT151 antibody. **C.** The summary statistics for barcoded subamplicon sequencing libraries. Total reads is the total number of paired-end sequencing reads obtained for that sample. Total aligned barcodes is the total number of barcodes (across all six subamplicons) that were aligned with at least two paired-end sequencing reads. Median effective depth is the median number of barcodes aligned per *env* codon.

We used deep sequencing to measure the frequency of each mutation in the mutant proviral plasmid, as well as in the mutant virus libraries with and without antibody selection. Because any given mutation is rare, we performed deep sequencing using a barcoded-subamplicon approach that tags individual molecules with unique molecular identifiers during the library preparation to increase accuracy [69–71]. Deep sequencing of the mutant proviral plasmids showed that 95% of the 12,559 possible single amino-acid mutations (661 mutagenized residues × 19 mutant amino acids) were present in each of the triplicate BF520 plasmid mutant libraries, with 99% present in the three libraries combined. Figure 2.3 shows that the sequencing error rate is 1.5×10^{-4} mutations/codon (equivalent to 5×10^{-5} mutations/nucleotide) as assessed by

sequencing unmutated wildtype proviral plasmid. This error rate is over 10-fold lower than the frequency of codon mutations in the plasmid mutant libraries.

We also quantified the rates of errors associated with viral replication by sequencing wildtype viruses passaged in parallel to each mutant virus library (described below). The rate of errors in these passaged wildtype viruses remained well below the rate of mutations in the virus mutant libraries, allowing us to reliably assess mutation frequencies (Figure 2.3). In the analyses below, we statistically correct for these errors associated with viral replication as described in the Methods.

Comprehensive profiling of escape from antibody PGT151

We chose to profile escape from the bnAb PGT151, which targets a quaternary, cleavage-dependent epitope made up of glycans and protein residues in the gp120/gp41 interface and fusion peptide. This bnAb has been extensively studied, using both structural [72,73] and functional approaches [25,74–76].

To generate mutant virus libraries, we transfected the BF520 mutant proviral plasmids into 293T cells and passaged the transfection supernatant in SupT1.CCR5 cells at a low MOI to establish a genotype-phenotype link and purge non-functional mutants. The resulting passaged mutant virus libraries, which we expect to carry virtually all Env amino-acid mutations compatible with viral replication, were used to infect cells in the presence and absence of PGT151. We then isolated and deep sequenced viral cDNA from infected cells to determine the frequency of each mutation, and computed the differential selection across Env.

Among the three biological replicates, the site differential selection was well correlated (R= 0.59-0.71; Figure 2.4A, Figure 2.5), indicating that the high-throughput experiments yielded reproducible results even when starting with fully independent mutant proviral plasmid libraries.

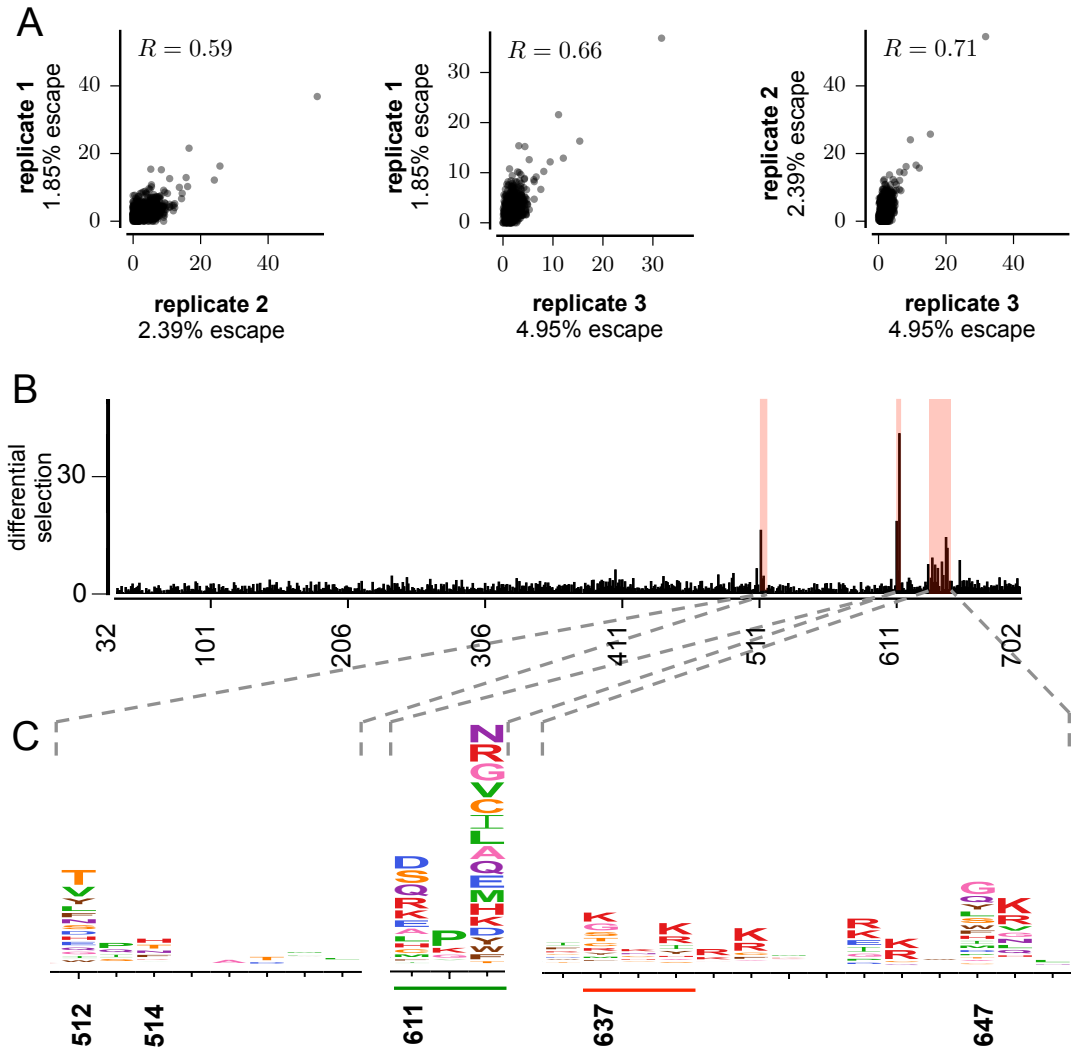


Figure 2.4. Reproducibility of mutational antigenic profiling.

A. Correlation of site differential selection across biological replicates. Each point indicates the selection at one of the 661 Env sites; plots also show the Pearson correlation. Percent escape was calculated by using droplet-digital PCR to measure the number of viral genomes in the antibody-

selected versus control sample. **B**. The site differential selection, averaged across replicates, is plotted across the Env sequence. See Figure S3 for data from individual replicates. **C**. Logoplots showing mutation differential selection in strongly selected regions. Mutations at numbered sites have previously been shown to reduce PGT151 neutralization sensitivity in other viral strains. From left to right: the fusion peptide, the 611 glycosylation motif (underlined in green), and the 637 glycosylation motif (underlined in red) and HR2 domain. Mutations can eliminate an N-linked glycan if they disrupt the N-X-S/T glycosylation motif, where X can be any amino acid except proline.

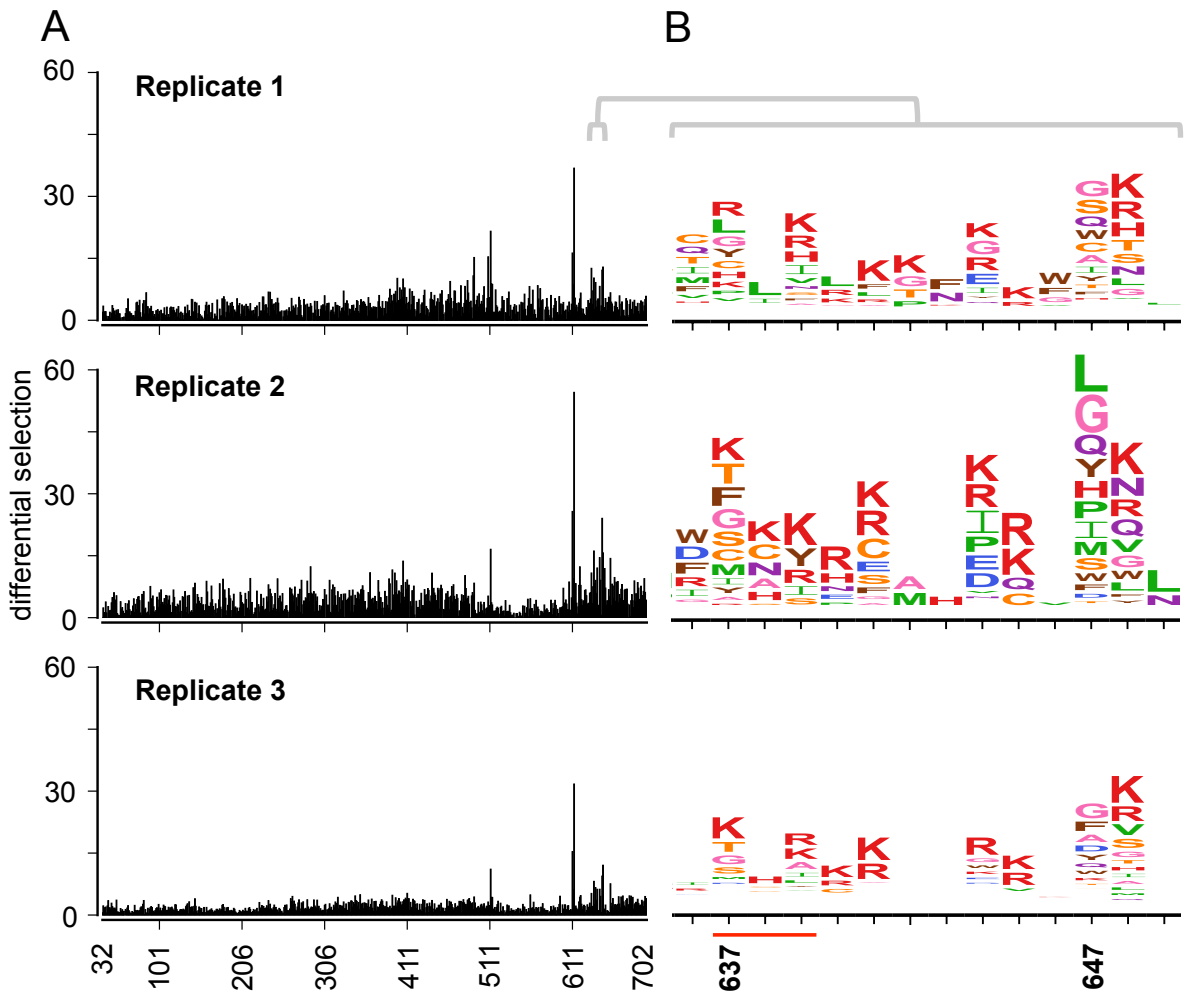


Figure 2.5. Reproducibility of mutational antigenic profiling. A. The site differential selection is plotted as in Figure 2.4C, but separately for each replicate. **B.** The escape profile of the 637 glycan motif (underlined in red) and HR2 domain is shown individually for each replicate. All logoplots are plotted on the same scale.

There was strong selection at sites in the known PGT151 epitope. Figure 2.4B plots the site differential selection across Env, and Figure 2.4C displays the escape profiles in regions of strong differential selection. These sites include residues originally mapped by mutagenesis and pseudovirus-neutralization assays: the dominantly targeted 611 glycan, the 637 glycan, and

residue 647 (Figure 2.4B,C) [74]. There was also strong differential selection at fusion-peptide sites 512 and 514, which have been mapped as part of the PGT151 epitope by structural [73] and functional [25,75,76] methods (Figure 2.4C). Therefore, our mutational antigenic profiling identified strong selection at epitope sites that have been identified by other approaches.

Validation of mutational antigenic profiling at previously mapped sites

We next used TZM-bl neutralization assays of BF520 pseudoviruses with select mutations that were and were not differentially selected in our mutational antigenic profiling to further examine how well our comprehensive escape profiles predicted neutralization escape (Figure 2.6). We first focused on sites that had been previously defined as conferring escape from PGT151, including mutations that disrupt the gp41 glycans at sites 611 and 637. Mutational antigenic profiling revealed strong selection for most mutations that disrupt the 611 glycosylation motif (N-X-S/T, where X can be any amino acid except proline) (Figure 2.6A). Both the 611 and 637 glycans make extensive contacts with the PGT151 Fab [73], and elimination of either can increase the IC_{50} in a number of strains [74]. Of note, there was selection for proline at the central position of the 611 glycosylation motif, but not at the central position in the 637 glycosylation motif (Figure 2.6A). The 638P mutation was depleted in the mutant virus libraries passaged in the absence of PGT151 selection, and we were unable to generate pseudoviruses bearing this mutation, suggesting that proline is not tolerated (data not shown).

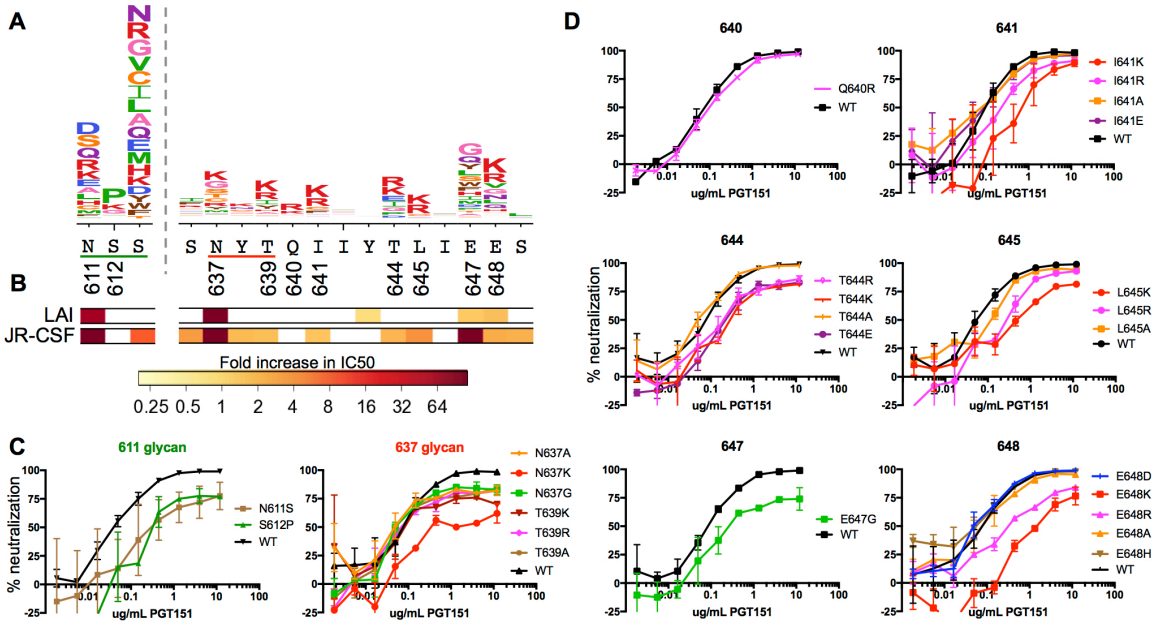


Figure 2.6. Analysis of the effect of individual amino acid changes identified by mutational antigenic profiling on PGT151 neutralization.

A. Logoplots showing differential selection. Sites where individual mutations were tested in neutralization assays are numbered, and the 611 and 637 glycosylation motifs are underlined in green and red. **B**. Color bars below the logoplots indicate the results of neutralization assays performed on point mutants (primarily to alanine) of JR-CSF and LAI by Falkowska et al. 2014. The bar at each site is colored according to the fold change in IC_{50} relative to wildtype that a mutation imparted. If multiple amino acids were tested at a site, the site is colored according to the largest effect. **C**. Results of TZM-bl pseudovirus neutralization assays on BF520 Env point mutants that alter the N-linked glycosylation motifs at sites 611 and 637. **D**. Results of neutralization assays at individual residues outside of glycosylation motifs. In panels C and D, each plot shows a representative neutralization curve.

At site 637, some of the glycan knockout mutations are under stronger differential selection than others . Consistent with the mutational antigenic profiling, the effect of these mutations on neutralization sensitivity in TZM-bl assays varied. According to the mutational antigenic profiling, the most strongly selected mutation at site 637 was to K .In the neutralization assays, this mutation increased the IC_{50} by 3.7-fold. Disrupting targeted glycans can also alter neutralization sensitivity by decreasing the maximum percent neutralization; the 637K variant also decreased the maximum percent neutralization to 59% (Figure 2.6C, 2.7A). All other tested mutations decreased the maximum percent neutralization of PGT151 to 74-81% (Figure 2.6C, 2.7A) but did not alter the IC_{50}. These results agree with and expand on previous observations; mutations that disrupt the 637 glycosylation motif have been shown to decrease the maximum percent neutralization in JR-CSF, LAI, and JR-FL [74], and PGT151 often incompletely neutralizes pseudoviruses bearing many different Env variants [77].

Prior work has also shown that mutations at site 647 reduce neutralization sensitivity in some strains (Figure 2.6B) [74]. Mutational antigenic profiling agreed with these results, revealing strong differential selection for many mutations at site 647. We confirmed the most strongly selected mutation (E647G) mediated escape in a neutralization assay, with IC_{50} fold-change of 5.3 (Figure 2.6D).

Overall, these results show that mutations at previously identified sites that are also mapped by our high-throughput approach indeed have clear effects on neutralization sensitivity in the BF520 variant when tested individually.

The comprehensive nature of mutational antigenic profiling allowed identification of new sites of escape

Our mutational antigenic profiling data showed strong selection at several sites where escape mutations have not previously been mapped. As noted above, both our work and prior studies identified escape mutations at site 647, where many different amino-acid mutations mediated escape (Figure 2.6A). In contrast, our mutational antigenic profiling also showed selection for mutations to positively charged amino acids at the neighboring site 648, where previous studies did not find effects of mutations to alanine or glycine (Figure 2.6A; [74]). Viruses engineered with E648K and E648R showed substantially reduced sensitivity to PGT151 (IC_{50} fold-change of 7.7 and 3.7, respectively), but mutations to a variety of other amino acids (D, A, and H) have no effect (less than 1.5-fold change in IC_{50}) (Figure 2.6D).

At a number of additional positions in the heptad repeat 2 domain (HR2), including sites 641, 644, and 645, we also observed strong differential selection for predominantly positively charged mutations. Consistent with prior work in other strains [74], both our mutational antigenic profiling (Figure 2.6A) and pseudovirus neutralization assays (Figure 2.6D) showed no effect of alanine mutations at these sites. However, the positively charged escape mutations revealed by mutational antigenic profiling indeed reduced neutralization sensitivity when assayed in individual neutralization assays, with fold changes in IC_{50} ranging from 3.2-5.7 fold.

Correlation between mutational antigenic profiling and traditional neutralization assays

At each validated site, the ranked order of effects of each mutation on neutralization sensitivity was well correlated between our high-throughput experiments and individual TZM-bl assays (Figure 2.6, 2.7). Across all validated sites, we next asked if our differential selection measures were representative of the true neutralization phenotype. For the BF520 mutations that we tested, the enrichment measured by mutational antigenic profiling is well correlated with the

fold-change in IC_{50} relative to wildtype from TZM-bl assays (R = 0.67, Figure 2.7). Remarkably, this is within the range of correlations observed between neutralization assays performed on the same sets of viruses by different laboratories [78]. Further, the IC_{50} only partially summarizes the information in a neutralization curve, as mutations also altered the maximum percent neutralization (Figure 2.7A). Many of the mutations that exhibit the poorest correlation between fold change in IC_{50} and enrichment in mutational antigenic profiling strongly affect the maximum percent neutralization (Figure 2.7A), suggesting that the differential selection measures captured both neutralization-curve phenotypes. Of note, the mutations that result in incomplete neutralization include those that eliminate targeted glycans as well as those that are in the Env-PGT151 protein interface. These data agree with observations that PGT151 often incompletely neutralizes *env* isolates, with PGT151 reaching less than 80% maximum percent neutralization for 26% viral isolates tested [74,77].

A

	Mutational antigenic profiling relative fold enrichment	TZM-bl neutralization	
		Fold increase in IC_{50} relative to WT	Maximum % neutralization
N611S	5.4	5.0	98
S612P	8.0	7.1	81
N637A	1.0	1.1	81
N637K	3.7	4.3	59
N637G	2.8	1.0	80
T639K	4.8	1.5	74
T639R	2.8	1.0	79
T639A	1.4	1.2	78
Q640R	2.6	1.1	98
I641K	5.7	5.8	91
I641R	3.2	2.6	92
I641A	1.4	1.3	101
I641E	1.2	1.4	98
T644R	4.0	2.8	87
T644K	3.2	5.3	80
T644A	1.0	1.1	96
T644E	2.2	2.3	84
L645K	3.8	7.6	90
L645R	3.7	3.3	93
L645A	0.7	1.8	96
E647G	5.3	2.6	74
E648D	0.7	0.9	98
E648K	7.7	7.8	74
E648R	3.7	10.9	104
E648A	1.2	1.5	101
E648H	1.5	1.0	113

<2	>90
2-4	80-90
4-6	<80
6-8	
>8	

B

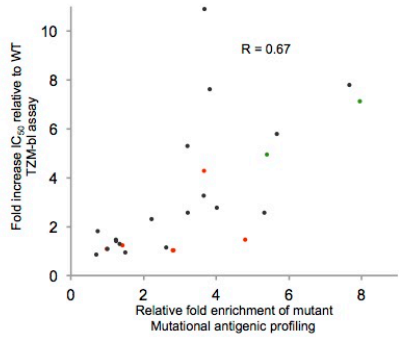


Figure 2.7. Comparison of effects of mutations in mutational antigenic profiling vs. TZM-bl neutralization assays. A. Comparisons between fold-enrichment in mutational antigenic profiling of each validated mutation (Figure 3), and the fold increase in IC_{50} value relative to wildtype and maximum percent neutralization estimated from replicate neutralization assays. **B**. For each individually tested mutations, the Pearson correlation between the fold-enrichment in mutational antigenic profiling and the neutralization assay fold-increase in IC_{50} relative to wildtype. Mutations that disrupt the 611 and 637 glycan are shown in green and red respectively.

Comparison of our escape profiles to the original gene-wide functional mapping of PGT151

The original residue-level mapping of the PGT151 epitope was based on neutralization of two large panels of pseudovirus mutants [74]. This work provided a large dataset against which we can benchmark our results and contrast our high-throughput measurements with results from traditional approaches. Figure 2.8A shows the differential selection that PGT151 exerts on every mutation at every site in Env with data from the original mapping of PGT151 shown in color bars below. In a single massively parallel experiment, we have recapitulated these original results (identifying sites 611, 637, and 647), identified sites mapped in subsequent studies (sites 512 and 514), and revealed new sites of escape in the HR2 domain. In contrast, examining mutations one-by-one in neutralization assays is exceedingly low throughput – for instance, even with herculean effort [74], traditional approaches have only managed to test mutations at a fraction of the sites in Env (Figure 2.8B). The discrepancy in completeness is even more apparent when considering the fact that testing one amino-acid mutation at a site does not reveal the effects of other mutations at the same site. Whereas our study examined all 12,559 amino-acid mutations, prior work has managed to examine only a small fraction of these mutations (Figure 2.8B).

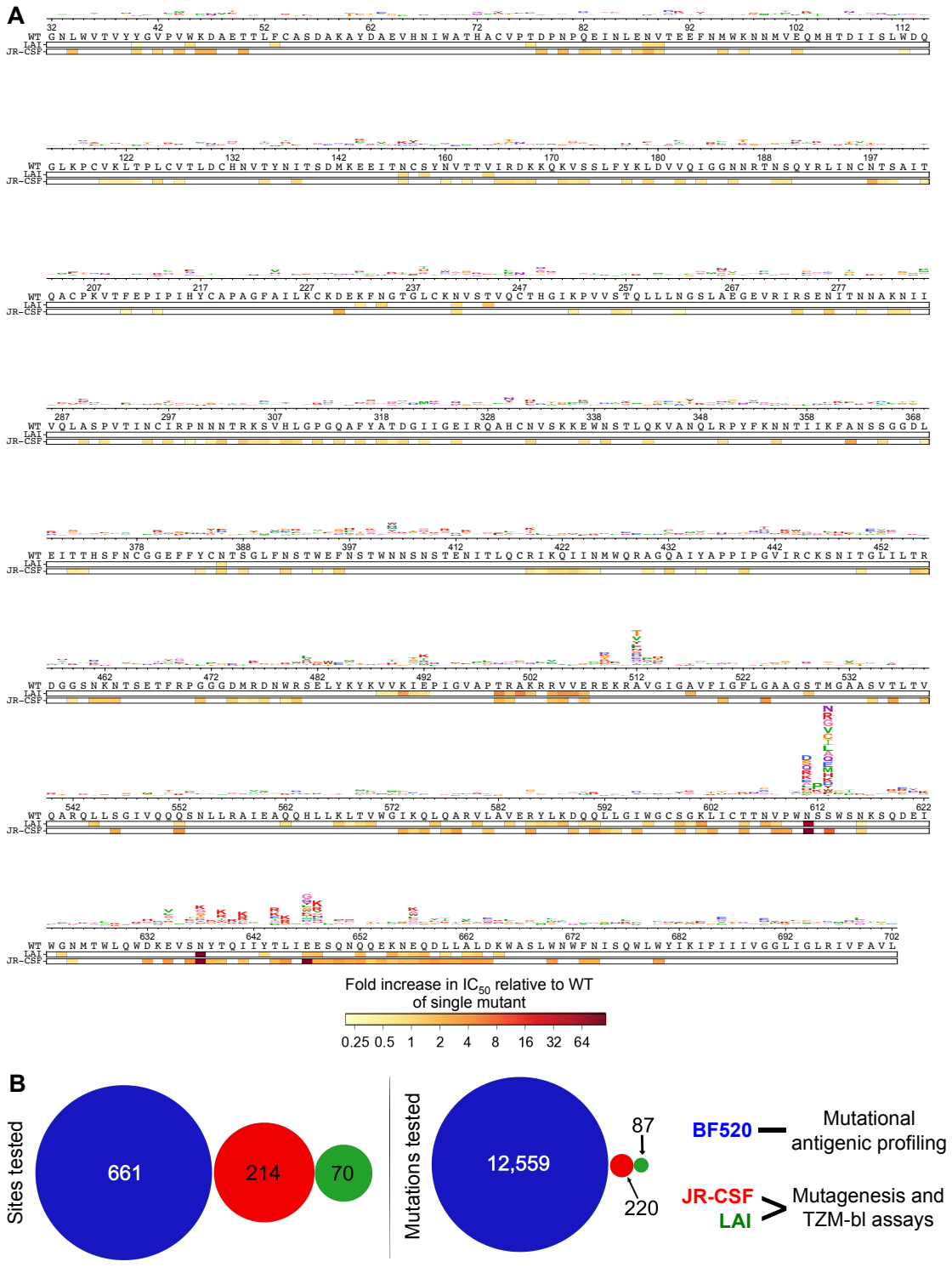


Figure 2.8. Complete mapping of escape from PGT151.

A. Differential selection for the entire mutagenized portion of Env. For systematic comparison, we plotted data from the original single residue-level mapping utilizing large panels of predominantly alanine scanning mutants of LAI and JR-CSF [74]; sites where mutations were tested are colored according to the fold increase in IC_{50} relative to WT as in Figure 2.6. The WT amino acid is also shown for each site. **B**. The number of sites and mutations tested for ability to replicate and escape PGT151 neutralization in the BF520 strain using mutational antigenic profiling compared with the number of sites and mutations tested using traditional approaches (mutagenesis and TZM-bl neutralization assays) applied to the JR-CSF and LAI strains in the original functional mapping of PGT151 [74].

Structural insights into biochemical basis of escape from PGT151

We can interpret the differential selection measured by mutational antigenic profiling in the context of known structural information about the interface between Env and PGT151. Figure 2.9A maps the maximum mutation differential selection at each site onto the cryo-EM model of JR-FL Env trimer bound by PGT151 Fabs [73]. The strongest differential selection maps to the structurally defined PGT151 epitope. The positively charged third complementary determining region of the heavy chain (CDRH3) of PGT151 extends into the negatively charged inter-protomer cavity at the interface between gp120 and the HR2 domain of gp41 (Figure 2.9). There is differential selection predominantly for mutations to positively charged amino acids on the CDRH3-proximal side of the HR2 alpha helix (Figure 2.9, Figure 2.10), suggesting that escape in this portion of the epitope is mediated by the introduction of charge-charge repulsions.

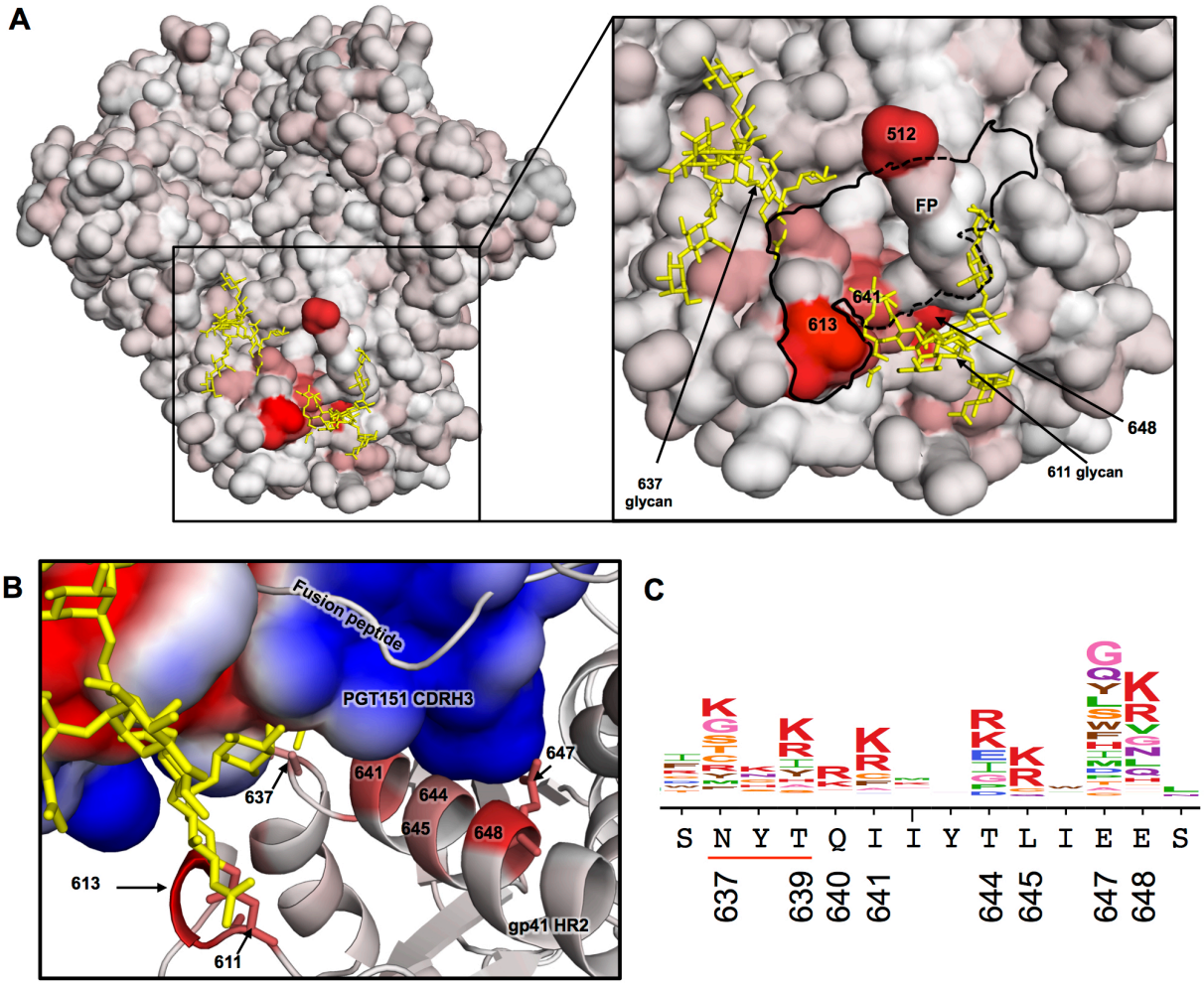


Figure 2.9. Mutational antigenic profiling combined with structural analysis suggests escape via the introduction of charge repulsions. A. The JR-FL EnvΔCT trimer cryo-EM model [73] with its surface colored according to the maximum mutation differential selection at each site. The 611 and 637 glycans are shown in yellow. The inset outlines the PGT151 footprint in black, defined as residues that come within 4 angstroms of the bound PGT151 Fab, with the 611 glycosylation motif residues included for clarity. The bound PGT151 Fab is not shown, and the fusion peptide (FP), which is sequestered into a hydrophobic pocket of bound PGT151, protrudes from the epitope in this view. **B**. A side view of the PGT151 CDRH3 – Env interface. Here, the ribbon representation of Env is colored according to the maximum mutation

differential selection at each site, while the surface of the bound PGT151 Fab is colored according to the Poisson-Boltzmann electrostatic surface potential (red to blue; negative to positive). The negatively charged side chain of site E647 is shown in stick representation. See Figure 2.10 for an alternate view of this interface. **C.** Logoplot showing the differential selection of the HR2 domain shown in B for reference.

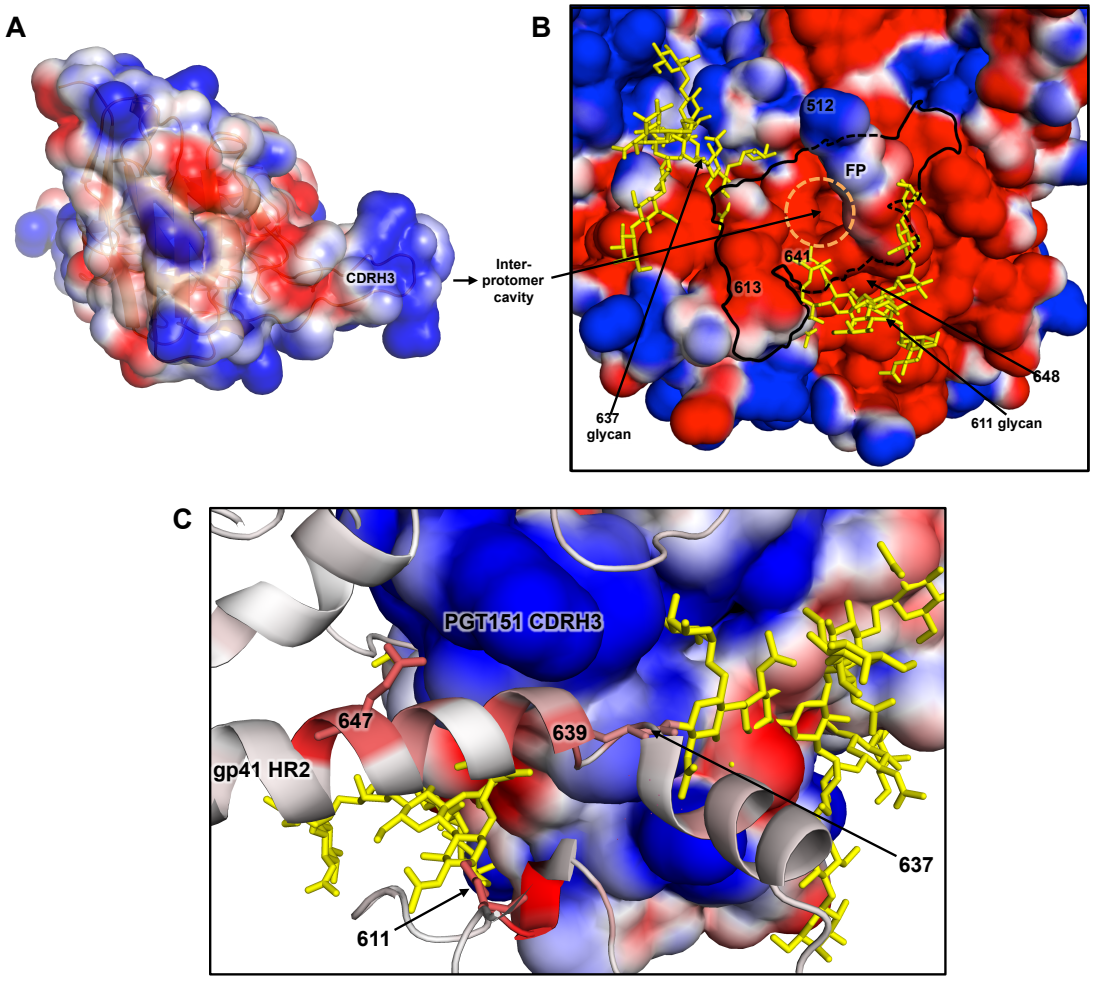


Figure 2.10. Electrostatic complementarity between PGT151 and the Env. A. The PGT151 Fab (in its trimer bound conformation) is colored according to the Poisson-Boltzmann electrostatic surface potential (red to blue; negative to positive). **B.** The surface representation of

Env, with the surface colored according to the electrostatic surface potential. The negatively charged inter-protomer cavity that the positively charged CDRH3 arm extends into is circled in orange. **C.** An internal view of the PGT151 CDRH3. Again, the ribbon representation of Env is colored according to the maximum mutation differential selection at each site, while the surface of the bound PGT151 Fab is colored according to the electrostatic surface potential.

The importance of charge-charge repulsions may extend to the glycosylated residue 637 (Figure 2.10). While any mutation that eliminates the 637 glycan results in incomplete neutralization, the introduction of a positively charged lysine results in a large reduction in both maximum percent neutralization and increase in IC_{50}. Similarly, mutating site 637 to lysine has a larger effect than mutating it to alanine in multiple other strains [74].

Indeed, it may be possible to conceptually distinguish between sites where almost any mutation mediates escape by *eliminating* key antibody-Env interactions, versus sites where only a few mutations mediate escape by *introducing* new steric or electrostatic clashes. For instance, site 647 appears to fall into the first category, with both our mutational antigenic profiling of BF520 (Figure 2.9C) and prior work with JR-CSF [74] suggesting that many mutations mediate escape. Indeed, the cryo-EM model suggests that the E at site 647 may interact with positively charged portions of PGT151's HCRD3 arm (Figure 2.9B, Figure 2.10C). In contrast, most of our newly identified sites of escape mutations, selection is for mutations to positively charged amino acids that are likely to clash with the positively charged antibody paratope (Figure 2.9B, Figure 2.10C). Such sites are especially difficult to identify by classical approaches such as alanine scanning, since the ability to escape neutralization is strongly dependent on which amino acid is introduced.

While the spatial clustering and biochemical evaluation of escape mutants suggests that they directly affect the interface between PGT151 and Env, it is possible that mutations indirectly affect neutralization by altering Env's conformation. To examine this possibility, we tested some of the most strongly selected PGT151 escape mutants with other bnAbs targeting the CD4 binding site and trimer apex. None of these mutations increased sensitivity to these other bnAbs targeting distal epitopes (data not shown). This fact strongly supports the idea that most PGT151 escape mutants that we have mapped directly disrupt the antibody-Env interface.

Discussion

We have developed a massively parallel approach to interrogate the neutralization phenotype of all functionally tolerated single amino-acid mutations to Env. We have used this approach to profile escape from PGT151, which targets a complex quaternary epitope. Our results recapitulate the escape mutations identified by previous low-throughput studies, reveal additional escape mutations not identified by these studies, and provide insight into the biochemical basis of escape.

Fine-resolution bnAb epitope mapping has previously been a cumbersome endeavor. Crystal structures of bnAb-Env complexes are often considered the gold standard, but can be difficult to obtain. In addition, the static structures do not reveal pathways of viral escape. Simply observing the evolution of HIV in the presence of a bnAb can identify specific escape mutations [36,38,39,44,45,50,79,80], but the stochasticity of evolution means that each study only identifies one of potentially many pathways of escape. Functional residue-level mapping traditionally relies upon generating and testing Env mutants one-by-one, a resource-intensive approach that can only be applied to a fraction of Env mutations. For instance, such studies

mostly restrict themselves to alanine mutations at surface sites, making them biased towards certain types of escape mutations. Therefore, despite the fact that PGT151 has been a subject of multiple studies using diverse techniques [25,72–76], our mutational antigenic profiling has yielded a far more comprehensive and unbiased map of viral escape than all this previous work.

This complete map offers a nuanced understanding of biochemical mechanisms of escape via evaluation of the physicochemical properties of the particular amino-acid mutations that enable escape at each site. For instance, our data suggest that introduction of electrostatic clashes between the CDRH3 of PGT151 and the HR2 domain of gp41 is a previously unappreciated mechanism of escape. The specificity of escape at these sites contrasts with the many diverse amino-acid mutations that were enriched at residue 647. This broad escape profile suggests that elimination of a key Env-PGT151 interaction underlies escape at site 647. The systematic quantification of the effect of each amino-acid mutation at each site has the potential to augment structural studies by providing insight into the energetics of antibody-epitope interactions.

The ability to comprehensively map viral escape also has important application for bnAb-based immunotherapies and vaccines. Quantifying how epitope features contribute to neutralization could aid in engineering broader and more potent antibodies [81]. Such information could also inform the design of immunogens that elicit responses that thwart common pathways of viral escape. Completely defining viral determinants of escape will also aid in evaluating the efficacy and failure of bnAb immunotherapeutics in humans, just as determining drug resistance mutations has aided the development of antiviral therapy and prophylaxis [82,83]. Similar to algorithms that leverage large datasets of drug-resistance mutations to predict antiviral resistance based on viral genotype [84], comprehensive mutation-level escape profiles could inform similar sequence-based scoring metrics for bnAb resistance.

As with all epitope-mapping approaches, mutational antigenic profiling has limitations. We are evaluating the effects of single amino-acid mutations to a single viral strain. Thus, this approach cannot reveal any potential synergistic effects of multiple mutations. Further, strain-specific differences in bnAb sensitivity have been described for PGT151 [72,74], and could explain why we did not observe escape at sites (533, 537, and 540) that have been shown to affect the sensitivity of JR-CSF to PGT151 [75]. Mutational antigenic profiling of escape from antibodies in multiple different Env isolates could determine the prevalence and mechanisms of such strain-specific differences. Lastly, we have assayed the enrichment of escape mutants at only a single strongly selective concentration of antibody. While our escape profiles accurately reflect values from TZM-bl neutralization assays, examining how mutation-level selection depends on antibody concentration is an interesting area for future work.

Few protein-protein interactions have been as heavily studied as those between bnAbs and Env. Indeed, these interactions provide the motivation for many current HIV treatment and prevention efforts. We have provided the first complete map of the viral determinants of neutralization and escape at one of these bnAb-Env interfaces. The mutational antigenic profiling approach that we have used to obtain this map is high-throughput and quantitative. We anticipate that this approach can be extended to define all possible HIV escape mutations from other bnAbs and possibly polyclonal serum. The resulting maps could be valuable for informing the design of immunogens and the development and evaluation of bnAb immunotherapies.

Methods

Generation of mutant DNA libraries

Codon mutant libraries were created in the context of BF520 *env* introduced into the full-length proviral clone Q23 [85]. Q23 is a subtype A transmitted/founder virus that is the basis for a well characterized system of generating chimeric full length proviral clones as well as pseudoviruses encoding heterologous *env* genes [86].

We first introduced codon mutations into *env* in three independent replicates using a slightly modified version of a previously PCR mutagenesis technique [87]. However, we made one modification: to overcome potential biases in the frequency of introduced mutations, we used mutagenesis primers of equal melting temperature (of ≈60°C) rather than equal length. We then cloned $4\times10^5 - 1\times10^6$ unique variants of for each mutant *env* library into a high efficiency-cloning vector (Q23.BsmBI.ΔEnv, File S5) that utilizes BsmBI restriction sites to clone *env* into the Q23 backbone. T4 DNA ligation products were transformed into ElectroMAX DH10B competent cells (Invitrogen; 12033-015). The next day, plated colonies were scraped, grown in LB plus ampicillin for 4 hours, and maxiprepped. The number of unique variants per library was determined by counting colonies on plates containing a dilution of the high efficiency transformation.

Generation of mutant virus libraries

To generate mutant virus libraries, 36 ug total of each mutant DNA library was transfected using Fugene-6 into 18 wells of 6-well plates, that had been plated with 4×10^5 293T cells per well the previous day. Two days post transfection, supernatants were collected, filtered through a 0.2 uM filter, and DNAse (Roche; 4716728001) treated to eliminate leftover transfection plasmid as previously described [51]. This transfection supernatant was then titered on TZM-bl cells by adding serial dilutions of the supernatant to 20,000 cells in the presence of

10 μg/mL DEAE-dextran in 600 μL total volume. After 48 hours, cells were fixed and stained for beta-galactosidase, and infected cell foci were counted.

To establish a genotype-phenotype link and select for functional variants, we passaged each mutant library for 4 days in SupT1.CCR5 cells at a low MOI (0.01 TZM-bl infectious units/cell). This initial passage in in SupT1.CCR5 cells was performed similarly to [51]. We infected cells in with 3×10^6 (replicate 1) or 4×10^6 (replicates 2 and 3) infectious units of transfection supernatant in T-225 flasks, each with 100 mL total volume of R10 (RPMI [GE Healthcare Life Sciences; SH30255.01], supplemented with 10% FBS, 1% 200 mM L-glutamine, and 1% of a solution of 10,000 units/mL penicillin and 10,000 μg/mL streptomycin) in the presence of 10 ug/mL DEAE-dextran with cells starting at a concentration of 1×10^6 cells/mL. On day 1, we replaced the media with fresh R10 containing 10 ug/mL DEAE-dextran, and on day 2, we doubled the total volume, splitting each flask into two. On day 4, we pooled all the flasks, spun down cells, and filtered the media through a 0.2 uM filter. We then concentrated this virus ~33 fold via ultracentrifugation for 1 hour at 4°C at 23,000 RPM over a 20% sucrose cushion using an SW 28 rotor (Beckman Coulter; 342207) and resuspended in R10. In parallel for each replicate, we passaged 5×10^5 infectious units of wildtype virus under the same conditions as a control. These passaged viruses were titered on TZM-bl cells.

PGT151 selection of mutant virus libraries

We incubated each mutant virus library with PGT151 and infected SupT1.CCR5 cells again. Each library was also passaged without PGT151 treatment to serve as a replicate-specific control to calculate differential selection. For each condition, 10^6 TZM-bl infectious units of each library was incubated +/- 1 ug/mL of PGT151 at 37°C for 1 hour, then infected into 10^6 (not

PGT151 treated) or 2×10^5 (PGT151 treated) SupT1.CCR5 cells in the presence of 100 ug/mL DEAE-dextran. This concentration was chosen with the goal of inhibiting ~97.5% of the viral infectivity. Three hours post infection, cells were spun down and resuspended in fresh R10, containing no DEAE-dextran. At 12 hours post infection, cells were spun down, washed with PBS, and then subjected to a mini-prep to harvest non-integrated viral cDNA. We also infected non-neutralized wildtype virus in parallel for each library.

Deep sequencing

To determine the frequency of each mutation in the antibody-selected and non-selected conditions, we utilized a barcoded subamplicon Illumina deep sequencing approach as previously described [51,69]. This approach uses unique molecular identifiers to distinguish true mutation from sequencing errors. It reduced the error rate when sequencing wildtype proviral plasmid to 1.5×10^{-4} mutations per codon (Figure 2.2). Figure 2.3C details the error-corrected sequencing depth per codon for each sample (roughly 10^5 unique error-corrected reads per codon for PGT151 selected samples).

Env Sequence numbering

Unless otherwise stated, Env residues are numbered according to the HXB2 reference strain numbering system [88]. The corresponding sequence numbering based on aligned BF520.W14M.C2 (GenBank accession number KX168094.1) and HXB2 env for the mutagenized portion of the gene is provided as File S3.

PGT151 production

PGT151 heavy (GenBank: KJ700282.1) and light (GenBank: KJ700290.1) chains were codon optimized, cloned into Igγ1 and Igκ expression vectors, and expressed in 293F cells using the FreeStyle MAX system (Invitrogen). IgG was purified using Protein G resin (Pierce; 20399) as described by Simonich et al., 2016.

ddPCR protocol

To quantify the remaining infectivity of each mutant virus library after PGT151 selection and infection into SupT1.CCR5 cells, we used droplet-digital PCR. The number of viral genomes present in each harvested sample of viral cDNA was quantified using a pol PCR [89] adapted for digital droplet detection [90]. The percent escape shown in Figure 2 was calculated using the number of genomes present in each selected library relative to its non-neutralized control.

TZM-bl neutralization assays

To validate our escape profiles, we generated pseudoviruses bearing individual point mutants of both enriched and non-enriched mutations. We then performed TZM-bl neutralization assays as previously described [91] in the presence of 10 ug/mL DEAE-dextran. Neutralization assays were performed in duplicate, and WT neutralization curves were run on each plate to reduce noise. Neutralization curves were fit using 3-parameter nonlinear regression of the % neutralization values across the dilution series, with the bottom plateau constrained to 0. To calculate the IC_{50} relative to wildtype, curves were solved for y = 50% neutralization. The fitted top plateau, averaged across replicates, is the maximum percent neutralization. Fold change in IC_{50} relative to WT was calculated within a single experiment and then averaged across replicates.

Structural analyses

All structural analyses are based on the cryo-EM model of JR-FL EnvΔCT trimer bound by two PGT151 Fabs ([73]; PDB ID: 5FUU). For all figures, we focused on PGT151-Env interface 2 [73]. Figures were generated using Pymol, and the Poisson-Boltzman electrostatic surface potential was calculated using the APBS plugin [92] with the protein dielectric constant of 20 for the trimer or PGT151 Fab (trimer bound conformation) in isolation.

Computation of differential selection

We calculated differential selection values as described in [61]. Briefly, the enrichment ($E_{r,x}$) of each amino acid x at site r relative to wildtype is calculated as shown in *equation 1*, where $n_{r,x}^{mock}$ is the number of counts of x at site r in the mock treated sample. Similarly, $n_{r,x}^{selected}$ is the number of counts of *x* at site r in the PGT151 selected sample. $wt(r)$ is the wildtype character at r.

$$E_{r,x} = \frac{\left(n_{r,x}^{selected} + f_{r,selected} \times P\right) / \left(n_{r,wt(r)}^{selected} + f_{r,selected} \times P\right)}{\left(n_{r,x}^{mock} + f_{r,mock} \times P\right) / \left(n_{r,wt(r)}^{mock} + f_{r,mock} \times P\right)} \qquad \textit{equation 1}$$

To account for statistical noise associated with low counts, a pseudocount of P =20 was added to each count. The $f_{r,selected}$ (*equation 2*) and $f_{r,mock}$ (*equation 3*) variables scale the pseudocount to account for different sequencing depth of the *mock* and *selected* libraries at site r.

$$f_{r,selected} = \max\left[1, \left(\sum_x n_{r,x}^{selected}\right) / \left(\sum_x n_{r,x}^{mock}\right)\right] \qquad \textit{equation 2}$$

$$f_{r,mock} = \max\left[1, \left(\sum_x n_{r,x}^{mock}\right) / \left(\sum_x n_{r,x}^{selected}\right)\right] \qquad \textit{equation 3}$$

To account for errors in viral replication and sequencing, counts in *equation 1* were adjusted by the rates of mutation to x at site r in mock selected wildtype virus libraries passaged in parallel

for each replicate. We define $n_{r,x}^{err}$ as the number of counts of x at site r in the matched wildtype virus control and $\epsilon_{r,x}$ as shown in *equation 4*, such that $\epsilon_{r,x}$ is the rate of errors to x at site r when $x \neq wt(r)$, and $\epsilon_{r,x}$ is one minus the rate of errors away from wildtype at site r when $x = wt(r)$ as defined in *equation 4*.

$$\epsilon_{r,x} = (n_{r,x}^{err})/(\sum_y n_{r,y}^{err}) \qquad \textit{equation 4}$$

The observed counts in *equation 1* are then adjusted to the error-corrected counts $\hat{n}_{r,x}$ as described in *equation 5*.

$$\hat{n}_{r,x} = \begin{cases} \max\left[\left(\sum_y n_{r,y}\right)\left(\frac{n_{r,x}}{\sum_y n_{r,y}} - \epsilon_{r,x}\right), 0\right] & if\ x\ \neq wt(r) \\ n_{r,x}/\epsilon_{r,x} & if\ x = wt(r) \end{cases} \qquad \textit{equation 5}$$

Lastly, differential selection values $s_{r,x}$ for x at *r* in the *selected* versus *mock* condition is quantified as $s_{r,x} = \log_2 E_{r,x}$, and visualized on logoplots rendered by dms_tools via weblogo [93]. Throughout this manuscript, we focused only on positively enriched mutations.

Data and Software Availability

Data and source code

Data were analyzed using dms_tools [94] version 1.1.20 as described in the iPython notebook computational pipeline available at https://github.com/adingens/BF520_MutationalAntigenicProfiling_PGT151. The software used to align sequence reads and compute differential selection is available at https://github.com/jbloomlab/dms_tools [94]. The deep sequencing data are available on the Sequence Read Archive under accession numbers SRX2548567- SRX2548579.

The script used to generate the codon tiling mutagenesis primers is available at https://github.com/jbloomlab/CodonTilingPrimers.

Chapter III

An antigenic atlas of HIV-1 escape from broadly neutralizing antibodies

The text in this chapter has been modified slightly from: bioRxiv 2018 Aug 31;
doi: doi.org/10.1101/406355, accepted at Immunity

Summary

Anti-HIV broadly neutralizing antibodies (bnAbs) have revealed vaccine targets on the virus's Env protein and are themselves promising immunotherapeutics. The efficacy of bnAb-based therapies and vaccines depends in part on how readily the virus can escape neutralization. While structural studies can define contacts between bnAbs and Env, only functional studies can define mutations that confer escape. Here we map how all single amino-acid mutations to Env affect neutralization of HIV by nine bnAbs targeting five epitopes. For most bnAbs, mutations at only a small fraction of structurally defined contact sites mediated escape, and most escape occurred at sites that are near but do not directly contact the antibody. The mutations selected by two pooled bnAbs were similar to those expected from the combination of the bnAbs' independent action. Overall, our mutation-level antigenic atlas provides a comprehensive dataset for understanding viral immune escape and refining therapies and vaccines.

Introduction

Over the last decade, a burgeoning number of broadly neutralizing antibodies (bnAbs) have been isolated from HIV-infected humans. These antibodies target conserved regions of Env that are promising vaccine targets [95]. Additionally, their broad neutralizing activity and

potential to direct the killing of infected cells make bnAbs promising antiviral immunotherapeutic drugs for HIV prevention, therapy, and cure strategies [40,96].

However, bnAbs face a formidable foe. HIV Env's exceptional evolutionary capacity allows the virus to stay one step ahead of bnAbs during infection, and resistance often arises when bnAbs are therapeutically administered to infected animals [38,39,49] or humans [43–47,50]. Thus, defining mutations that mediate viral escape is essential to optimizing and evaluating bnAb immunotherapies and vaccines.

While extensive efforts have gone into structurally characterizing bnAb epitopes via X-ray crystallography and cryo-electron microscopy (cryo-EM), structures on their own are insufficient to completely define the *functional epitope* [57,97], defined here as sites where mutations affect antibody neutralization of replication-competent virus. Making individual mutations to Env and performing neutralization assays can provide information on the functional effect of specific mutations, but even the largest studies employing one-at-a-time mutagenesis can only assay a small fraction of all possible Env mutations.

We recently described mutational antigenic profiling, a massively parallel experimental approach to quantify how all single amino-acid mutations to Env affect antibody neutralization [53]. This approach involves generating libraries of HIV that carry all Env amino-acid mutations compatible with viral replication, incubating these libraries with or without an antibody, infecting T cells, and using deep sequencing to quantify the enrichment of each mutation in the selected versus non-selected libraries. Here, we apply this approach to a panel of nine bnAbs that target five Env epitopes, as well as a pool of two bnAbs. The resulting maps of viral escape provide comprehensive mutation-level views of the functional interfaces between HIV and bnAbs.

Results

Complete maps of viral escape from a panel of bnAbs

To gain a broad picture of viral escape, we selected nine bnAbs targeting the five best-characterized epitopes on Env (Figure 3.1A). Specifically, the bnAb panel includes the CD4 binding site (CD4bs) bnAbs VRC01 [98] and 3BNC117 [99], the V3/N332 glycan supersite bnAbs PGT121 [100] and 10-1074 [101], the V2 glycan/trimer apex bnAbs PG9 [102] and PGT145 [100], the fusion peptide and gp120/gp41 interface bnAbs PGT151 [74] and N123-VRC34.01 (subsequently referred to as VRC34.01, Kong et al., 2016), and the membrane proximal external region (MPER) bnAb 10E8 [103]. The binding footprints of these antibodies have been previously characterized using structural techniques (Figure 3.1B), allowing us to compare the *structural* epitopes with the *functional* epitopes defined by this study.

We mapped escape from these antibodies using the BG505.T332N Env, which is from a transmitted-founder subtype A HIV strain [63]. This Env trimer is used widely in structural and vaccination studies (Sanders et al., 2013; Ward and Wilson, 2017). We used viral libraries that were previously generated by making all possible amino-acid mutations to the ectodomain and transmembrane domain of Env. There are 19 amino-acid mutations × 670 sites = 12,730 such amino-acid mutations, and our libraries contain the subset of these mutations that is compatible with viral growth in cell culture [51,52].

To quantify how each of these mutations affect HIV's antibody sensitivity, we neutralized independently generated mutant virus libraries at an ~IC_{95}-$IC_{99.9}$ antibody concentration, and deep sequenced the *env* genes of viruses that were able to infect cells in the presence of antibody (Figure 3.1C). For each antibody we performed at least two replicates using

independently generated viral libraries. We performed parallel control experiments without antibody, and compared the relative frequency of each Env mutation in the antibody-selected library to the non-selected control. By scaling this relative frequency by the overall fraction of the entire library that survived the antibody selection, we estimated the fraction of virions with each mutation that survive the selection, hereafter termed the *fraction surviving* [106]. To highlight escape mutations, we plotted the excess fraction surviving above the overall library average.

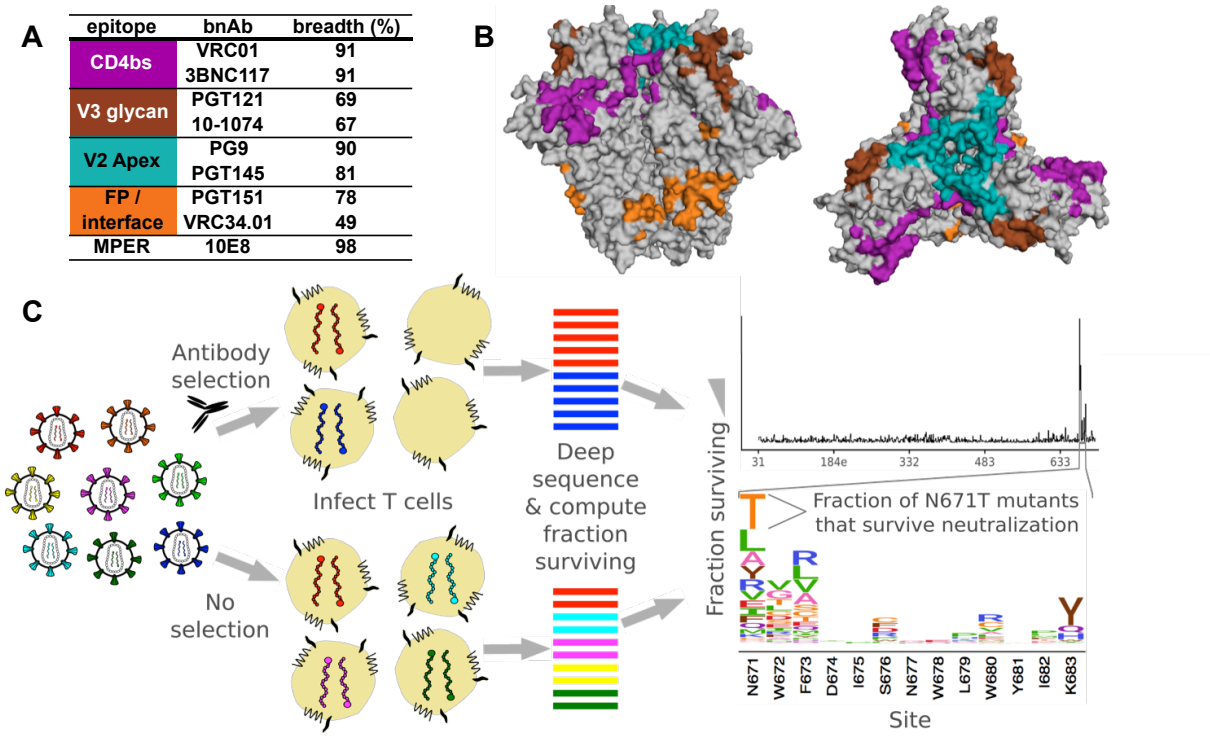


Figure 3.1. Schematic of mutational antigenic profiling of a panel of bnAbs. A. The bnAb panel. Breadth measures are from the 100 most commonly used viruses on LANL's CATNAP [107]. **B.** For each epitope, structurally defined antibody contact sites are indicated by colors on the side and top view of the BG505 SOSIP Env trimer (PDB:5FYL). **C.** The mutational antigenic profiling experimental workflow, and example data from bnAb 10E8. A mutant virus library is incubated with or without and antibody before infecting SupT1.CCR5 cells. Non-integrated viral

cDNA from infected cells is deep sequenced to quantify the frequency of each *env* mutation in both the antibody-selected and mock-selected conditions, and the overall fraction of the virus library that survives antibody neutralization is quantified via qPCR. The fraction of each mutant that survives neutralization is plotted at the site level in line plots, and at the mutation level in -logoplots. The height of each letter is proportional the fraction of the virions with that amino acid that survived antibody selection in excess of the overall library average.

Each antibody reproducibly selected mutations at just a small subset of Env sites (Figure 3.2A, Figure 3.3). Antibodies targeting the same epitope tended to select mutations in similar regions of Env, and these mutations cluster in three-dimensional structure in or near the antibody-binding footprint (Figure 3.2A). This is clearly exemplified by PG9 and PGT145, where the selected positions largely overlapped. Note also that the effect size of mutations varied across antibodies (compare the y-axes in Figure 3.2A) as did the apparent noise in the plots; the implications of this are discussed in more detail in a later Results subsection.

To rigorously compare the overlap between structural contacts and sites of viral escape, we identified sites of significant functional escape from the mutational antigenic profiling data, and sites of physical contact between the bnAb and Env from published structures. We defined significant sites of viral escape by fitting a gamma distribution to the measured antigenic effects of mutations at each site, and identified sites where the antigenic effects were larger than expected from this distribution at a false discovery rate of 0.01 (Figure 3.4). We defined sites of structural contact as Env residues that were within 4 Å of the antibody in available structural models, only considering non-hydrogen atoms (see Methods for details).

For most antibodies, only a small fraction of the structurally defined contact sites were also sites of significant viral escape (Figure 3.2B). Further, we identified numerous sites of escape outside the structurally defined epitope for most antibodies. The extent to which escape occurred at sites that directly contact the antibody differs considerably across bnAbs, ranging from all significant sites of escape for PG9 to only one of five sites of escape for VRC01 (Figure 3.2C). The sites of escape that do not directly contact the antibody are usually near the structurally defined epitope, in the 5-10 Å range. However, a few sites of escape are more distant from the structurally defined epitope (Figure 3.2A).

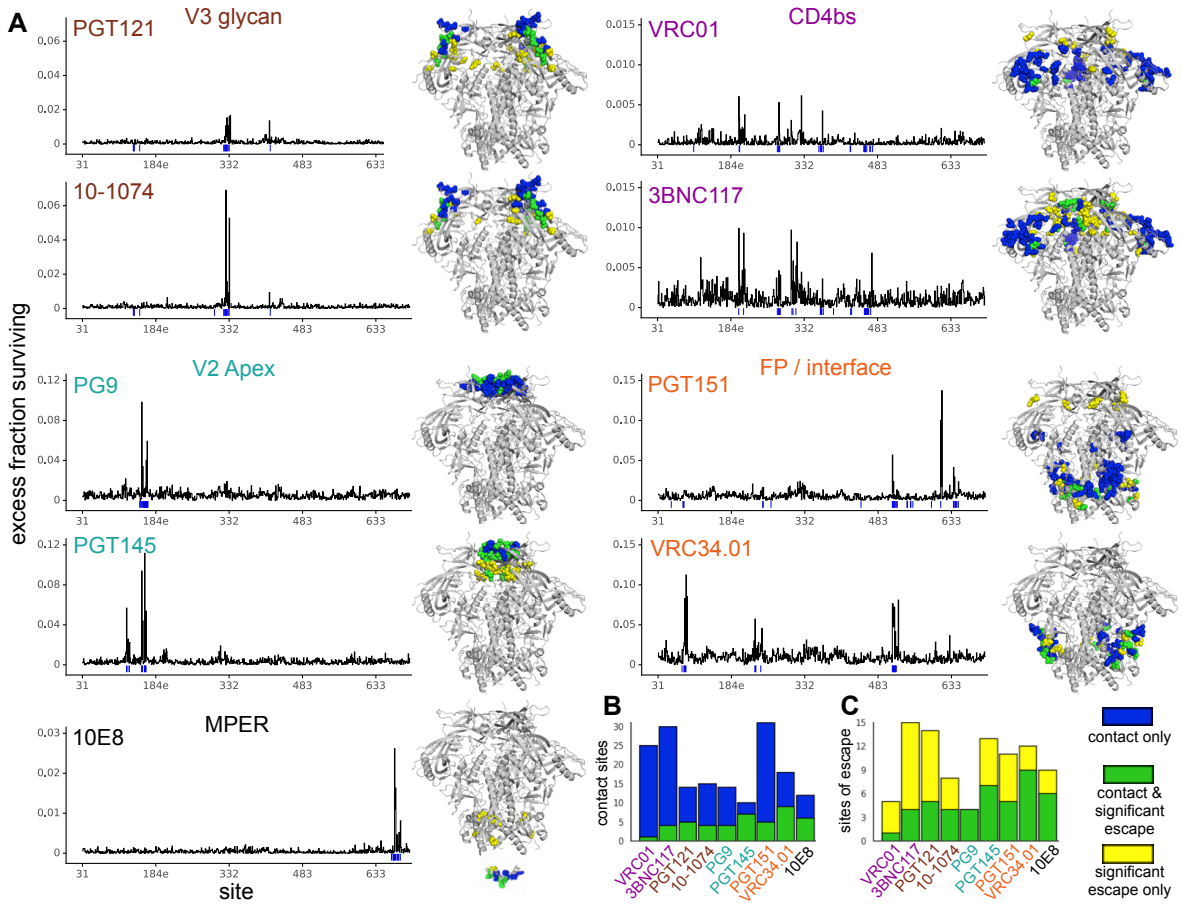


Figure 3.2. Env-wide escape profiles in relation to amino-acid positions that contact the bnAb. A. The line plots show the excess fraction surviving antibody neutralization averaged

across all mutations at each site. Structurally defined contact sites are indicated by a blue line. The structures show the BG505 Env SOSIP trimer (PDB:5FYL) with sites of significant escape colored yellow, contact sites colored blue, and overlap between these sets of sites colored green. For 10E8, the MPER peptide structure (which is absent from the SOSIP trimer) is also shown (PDB:4G6F). **B.** Bars give the number of structurally defined contact sites for each antibody, with green indicating the contact sites that are also sites of significant escape. **C.** Bars give the number of sites of significant escape for each antibody, with green indicating the sites of escape that are also contact sites. Note that the green bars encompass the same sets of sites in panels B and C.

A

antibody	antibody concentration per mL	library	fraction surviving	percent neutralized
VRC01	11ug	1	0.002029	99.80
	8ug	2	0.027775	97.22
	11ug	2	0.009827	99.02
	11ug	3	0.001008	99.90
	11ug	3	0.007226	99.28
3BNC117	4ug	1b	0.008765	99.12
	3ug	2b	0.009847	99.02
	4ug	3b	0.006475	99.35
	4ug	2b	0.022027	97.80
PGT121	4-25ug	2b	0.003383	99.66
	3-5ug	2b	0.007412	99.26
	3-5ug	3b	0.012455	98.75
	3-5ug	1	0.010645	98.94
10-1074	3ug	2b	0.019202	98.08
	4ug	2b	0.015257	98.47
	4ug	3b	0.003771	99.62
	4ug	1b	0.011481	98.85
PG9	10ug	1	0.025034	97.50
	10ug	3	0.016036	98.40
PGT145	125ug	2b	0.028826	97.12
	125ug	1b	0.031111	96.89
	125ug	3b	0.033075	96.69
PGT151	0.2ug	3	0.035396	96.46
	0.2ug	1	0.020718	97.93
VRC34	33ug	3	0.050239	94.98
	33ug	1	0.063644	93.64
10E8	60ug	1	0.001157	99.88
	60ug	3	0.002491	99.75

B

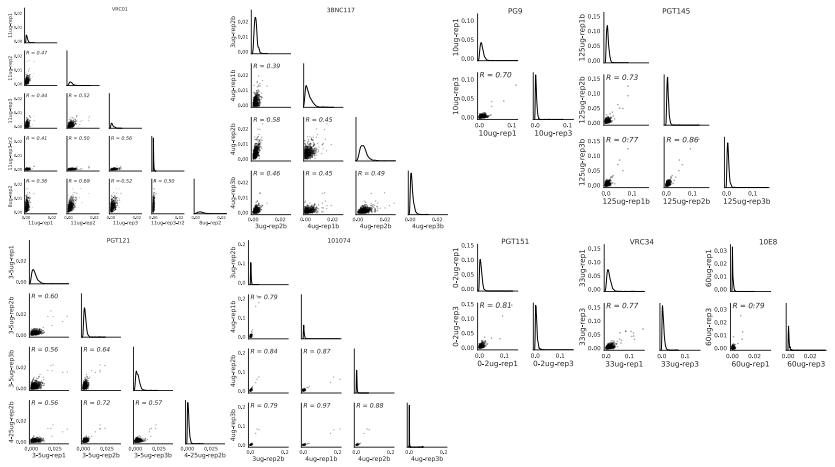
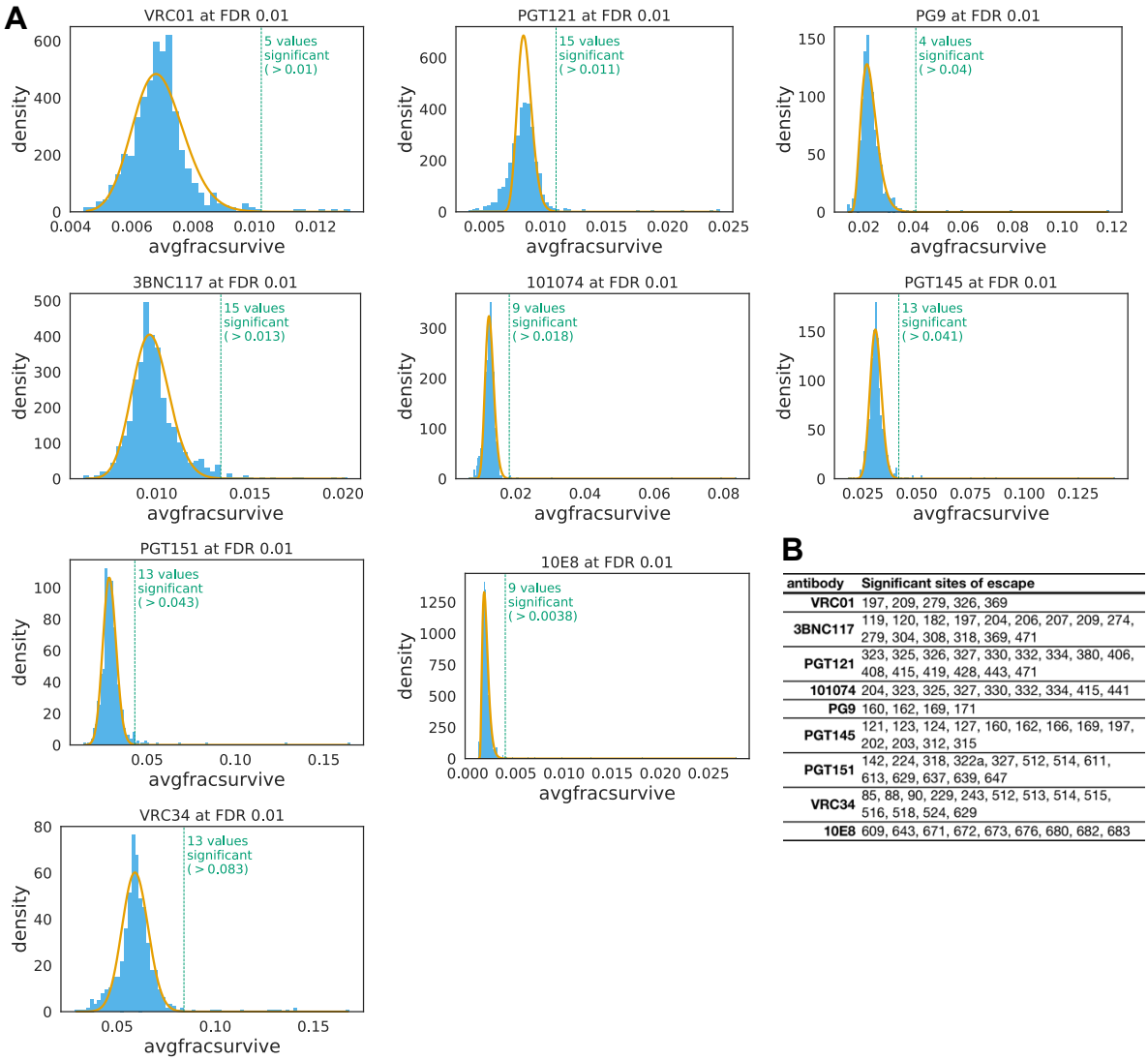


Figure 3.3. The fraction surviving measurements and correlation between mutational antigenic profiling biological replicates. A. For each biological replicate, the antibody concentration used during the selection, which mutant virus library was used, and the fraction of that library that survived antibody selection is shown. For clarity, the percent neutralized (1-library fraction surviving) × 100 is also shown. **B.** The correlation between the average excess fraction surviving at each site for each biological replicate, for each antibody.



B

antibody	Significant sites of escape
VRC01	197, 209, 279, 326, 369
3BNC117	119, 120, 182, 197, 204, 206, 207, 209, 274, 279, 304, 308, 318, 369, 471
PGT121	323, 325, 326, 327, 330, 332, 334, 380, 406, 408, 415, 419, 428, 443, 471
101074	204, 323, 325, 327, 330, 332, 334, 415, 441
PG9	160, 162, 169, 171
PGT145	121, 123, 124, 127, 160, 162, 166, 169, 197, 202, 203, 312, 315
PGT151	142, 224, 318, 322a, 327, 512, 514, 611, 613, 629, 637, 639, 647
VRC34	85, 88, 90, 229, 243, 512, 513, 514, 515, 516, 518, 524, 629
10E8	609, 643, 671, 672, 673, 676, 680, 682, 683

Figure 3.4. Identification of significant sites of viral escape. A. For each antibody, the distribution of the average fraction surviving at each site is plotted in blue. A gamma distribution fit to the site fraction surviving values using robust regression is overlaid in yellow. A dotted line shows sites that fall beyond this distribution at a fall discovery rate of 0.01, and the number of sites that beyond this cutoff is labeled. Code that performs this analysis is at https://jbloomlab.github.io/dms_tools2/dms_tools2.plot.html#dms_tools2.plot.findSigSel. **B.** A table listing all of the significant sites of viral escape for each antibody.

A

10-1074

Mutation	Mutation fraction surviving	Fold Change IC_{50} relative to WT	Mutational effect on viral growth
C119R	0.008278	0.8	0.097958
D140E	0.008363	1.6	1.533243
R151K	0.011343	1.6	0.767443
D322A	0.009009	0.8	1.117847
D322E	0.008749	1.4	0.834735
D322G	0.006662	1.4	0.889122
I323G	0.035657	2.5	1.472166
D325E	0.063118	27.2	1.383803
D325N	0.010407	0.7	1.410747
D325S	0.044115	3.0	2.545381
I326Y	0.009754	0.9	0.42606
R327A	0.035596	4.0	0.527711
H330R	0.048421	35.1	0.114338
T415R	0.068193	4.8	1.050657
T415Y	0.041935	2.7	1.76969
G441P	0.014016	2.2	0.24461

PGT121

Mutation	Mutation fraction surviving	Fold Change IC_{50} relative to WT	Mutational effect on viral growth
C119R	0.01423	1.2	0.097958
D140E	0.002981	1.6	1.533243
R151K	0.008765	1.3	0.767443
D322A	0.018474	1.1	1.117847
D322E	0.004105	1.4	0.834735
D322G	0.009586	1.4	0.889122
I323G	0.032053	1.9	1.472166
D325E	0.009864	1.7	1.383803
D325N	0.003128	0.8	1.410747
D325S	0.007069	1.2	2.545381
I326Y	0.008386	1.3	0.42606
R327A	0.037228	3.4	0.527711
H330R	0.017344	1.6	0.114338
T415R	0.073432	3.8	1.050657
T415Y	0.034631	2.4	1.76969
G441P	0.008918	2.6	0.24461

VRC01

Mutation	Mutation fraction surviving	Fold Change IC_{50} relative to WT	Mutational effect on viral growth
D113N	0.014156	4.1	0.040187
C119R	0.011311	1.6	0.097958
D140E	0.008087	1.8	1.533243
R151K	0.015409	1.7	0.767443
N197E	0.005161	0.02	0.272265
N197S	0.058665	26.1	0.380752
T198A	0.039597	3.6	0.15237
N279K	0.035774	>175	0.006956
I323G	0.005812	1.1	1.472166
D325E	0.002185	0.8	1.383803
D325N	0.002956	0.8	1.410747
D325S	0.005692	0.9	2.545381
I326Y	0.042948	2.9	0.42606
R327A	0.010742	1.3	0.527711
H330R	0.009456	0.9	0.114338
L369K	0.031735	4.01	0.888277
T415R	0.007226	0.9	1.050657
T415Y	0.00695	0.9	1.76969
G441P	0.011296	1.4	0.24461

PGT145 Data from Lee *et al. Immunity* 2017

Mutation	Mutation fraction surviving	Fold Change IC50 relative to WT	Mutational effect on viral growth
K121A	0.069501	302	0.18497
V127A	0.049786	302	0.065149
N156K	0.075847	302	0.029769
N160A	0.061076	302	0.175816
N160K	0.091395	302	0.410596
M161A	0.022357	60.4	1.083304
T162A	0.112894	302	0.112595
E164A	0.041677	2.7	0.056266
L165A	0.031111	167.1	0.006771
R166A	0.196633	302	1.510136
D167A	0.031111	302	0.01187
K168A	0.025926	274.8	0.091057
K169A	0.046097	256.2	0.80713
I307A	0.028948	0.3	0.005547
I420A	0.020672	3.8	0.949832
I423A	0.033075	1.9	0.014982

B

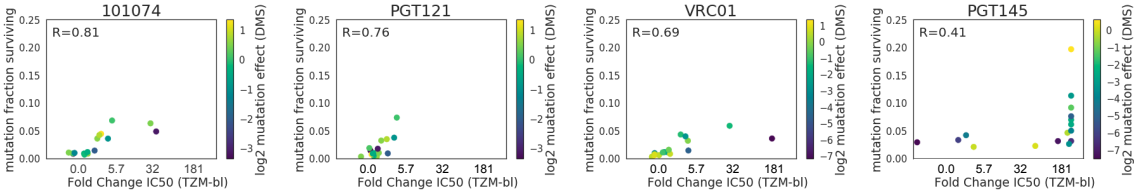


Figure 3.5. Validation of mutational antigenic profiling results using BG505 point mutants in TZM-bl neutralization assays. A. The tables list mutations that were selected for validation in neutralization assays. Data for PGT145 was taken from [20]. For each mutation, the table gives the mutation's fraction surviving antibody, fold change in IC_{50} in the TZM-bl neutralization assay, and the mutation's effect on viral growth. The mutation's effect on viral growth is calculated from our prior deep mutational scanning of Env for viral growth in cell culture, in the absence of any immune selection [52]. The mutational effect is the ratio of the preference for that mutant amino acid relative to the wildtype amino acid at that site. If the log mutational effect is >0, then that mutation grows better than wildtype in cell culture; if it is <0, that mutant grows worse than wildtype. **B.** For each antibody, we plot the correlation between the mutation fraction surviving and the fold change in IC_{50} relative to wildtype from TZM-bl assays. Points are colored according the $\log_2$ mutational effect of the mutation on viral growth. Where there are discrepancies between mutational antigenic profiling and TZM-bl assays, the mutational antigenic profiling estimate is usually less than would be predicted by the TZM-bl

assay. These mutants (for example, N279K for VRC01 and H330R for 10-1074) often have a $\log_2$ mutational effects on viral growth <<0 (darker blue), indicating they are deleterious for viral growth in cell culture. Of note, mutating site 279 has also been shown to have a fitness cost in other assays [80]. Thus, the mutational antigenic profiling datasets may underestimate the antigenic effect of high fitness cost mutations, though these measures may possibly better reflect the ability of a mutant to both replicate and escape antibody neutralization *in vivo* than TZM-bl point mutant neutralization sensitivities alone. It is important to note that determining an exact cut-off between biologically relevant signal and noise is difficult in both traditional TZM-bl neutralization assays and mutational antigenic profiling.

While our maps of escape include most mutations previously identified using individual BG505 point mutant pseudoviruses in TZM-bl neutralization assays [20,25], we also uncovered many previously uncharacterized sites of escape. We generated and tested BG505 point mutant pseudoviruses in TZM-bl neutralization assays for three antibodies, testing 16 to 19 point mutants for each antibody. The measurements from the mutational antigenic profiling were well correlated with the fold change in IC_{50} from TZM-bl neutralization assays for all antibodies tested (PGT121: R=0.76, n=16; 10-1074: R=0.81, n=16; VRC01: R=0.69, n=19) (Figure 3.5). In the next few subsections, we focus on each Env epitope individually.

Escape from V3 glycan supersite bnAbs

The two anti-V3 antibodies PGT121 and 10-1074 are clonal variants that arose in the same infected individual [101]. However, there were intriguing differences between the two

antibodies in the specific mutations that mediated escape in our experiments, as well as the overall effect sizes of escape mutations (Figure 3.6A,B, note the different y-axis scales in the two panels). For instance, mutations to site 325 had a larger effect on 10-1074 than PGT121, whereas mutations at site 327 had similar effects. We validated the differential effects of mutations to site 325 by testing three different mutants at this site in TZM-bl neutralization assays: the maximal effect for 10-1074 was a 27-fold increase in IC_{50}, while the maximal effect for PGT121 was just a 1.7-fold increase (D325E; Figure 3.5). Our mutational antigenic profiling also indicated that mutations that eliminated the N332 glycan had a larger effect for 10-1074 than PGT121 (Figure 3.6A,B), consistent with a prior study that examined binding to gp120 [101]. In contrast, at most other sites in the epitope (such as 323, 327, and 330), the overall effects of mutations were similar between the two antibodies (Figure 3.6A,B).

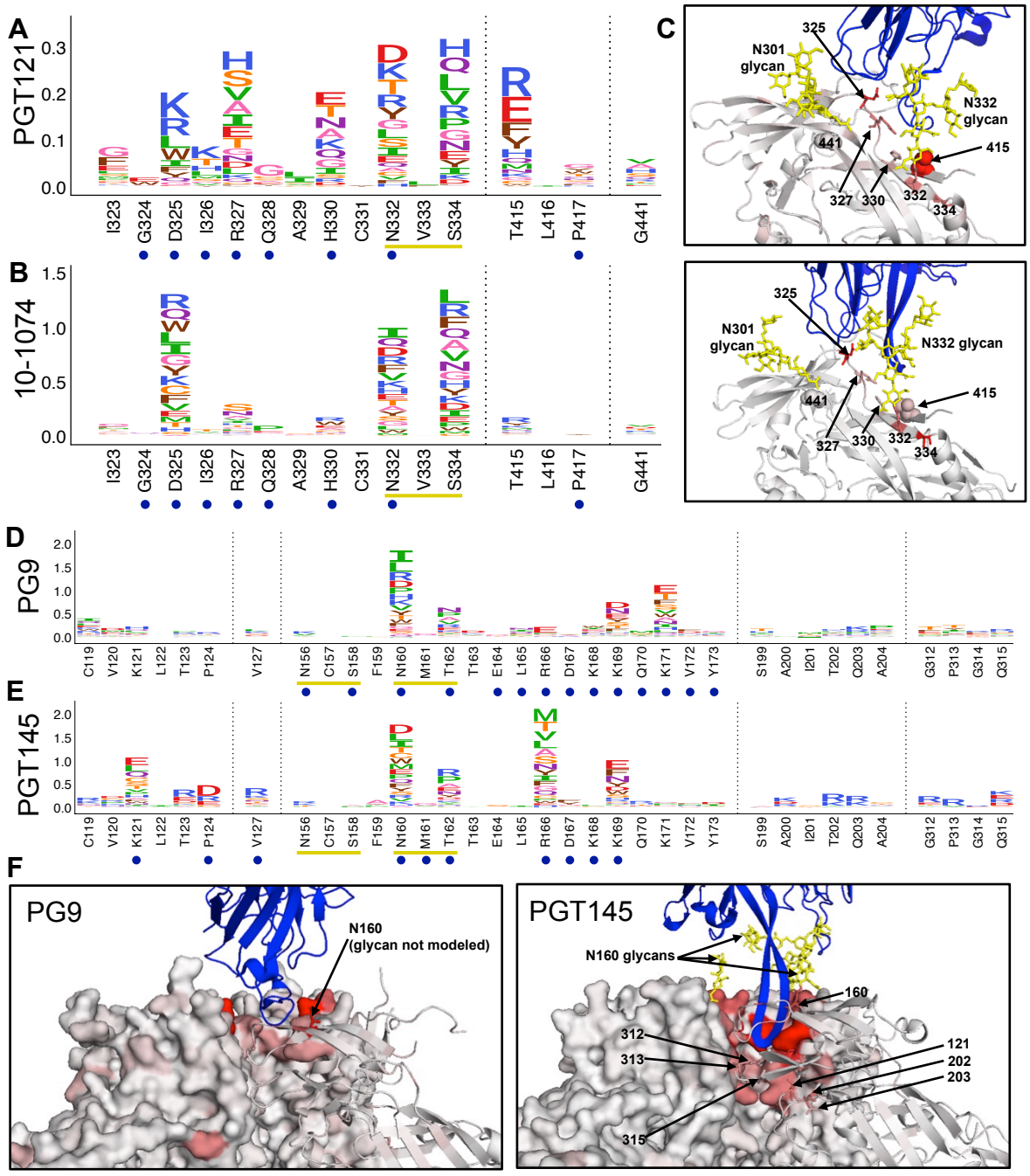


Figure 3.6. Escape From V3 glycan supersite and V2 apex bnAbs. A, **B.** Escape profiles for V3 glycan supersite bnAbs PGT121 and 10-1074. Letter heights indicate the excess fraction surviving for each mutation. Blue circles indicate structurally defined contact sites, and yellow underlines indicate a N-linked glycosylation motif. **C.** V3 glycan supersite antibodies are shown in blue, and Env is colored according to the maximum excess fraction surviving at each site.

Note that for PGT121, the closely related clonal variant PGT122 structure is used in lieu of a PGT121 structure (PDBs: 5FYL and 5T3Z respectively). **D**, **E.** Escape profiles of V2 glycan/apex antibodies PG9 and PGT145, presented in the same manner as A, B. **F.** V2 glycan/apex antibodies are shown in blue, and Env is colored according to the maximum excess fraction surviving at each site (PDBs: 5VJ6 and 5V8L respectively).

There are also differences in *which* mutations at a given site escape each antibody. For example, while the overall effect of all mutations at site 330 is similar between 10-1074 and PGT121, H330R escapes 10-1074 but not PGT121 (Figure 3.6A,B, validated by TZM-bl neutralization assays in Figure 3.5). Even among small-effect mutations, TZM-bl neutralization assays validated the results of mutational antigenic profiling. For example, mutations at site 325 had disparate effects: D325E had a large effect on 10-1074 but a negligible effect PGT121, D325S had a small effect for 10-1074 but no effect for PGT121, and D325N had no effect on either antibody (Figure 3.6A,B and validated in Figure 3.5).

Many aspects of our maps of escape are consistent with prior knowledge about the epitopes of these two antibodies [101,108–110]. For example, eliminating the targeted N332 glycan via mutations to N332 and S334 resulted in escape from both antibodies, as did antibody-specific mutations in the ^{324}GDIR^{327} motif (Figure 3.6), a conserved region of this epitope that is involved with CCR5 co-receptor binding [108].

However, we also identified escape at sites not previously implicated as being part of the functional epitope. For instance, viral escape from both antibodies occurred via mutations at site 415 in V4, and to a modest but reproducible extent, site 441 (Figure 3.6A, B). We validated that mutations at each of these sites resulted in escape from both antibodies in TZM-bl neutralization

assays (Figure 3.5). Neither of these sites directly interact with either antibody; site 415 it is close to other structural contacts as well as the N332 glycan, and site 441 in the β22 strand neighbors the N301 glycan.

Escape from V2 apex bnAbs

For the V2 apex antibodies PG9 and PGT145, escape occurred via eliminating the N160 glycan at the heart of the epitope (Figure 3.6D,E). Additionally, escape occurred at structurally defined contact sites at the trimer apex for PG9, and at the trimer apex and interface for PGT145. For both antibodies, prior studies suggest that binding is driven by electrostatic interactions with positively charged Env residues [20,111,112]. Mutations to these sites resulted in viral escape (including residues R166, K169, and K171 for PG9, and K121, R161, K169 for PGT145), with charge swaps often resulting in the greatest extent of escape (Figure 3.6D,E). Structural studies indicate that the long HCDR3 arm of PGT145 reaches into the trimer interface; existence of this epitope has been hypothesized to result from a balancing act of a "push" from inter-protomer charge repulsions at the trimer interface and a "pull" of hydrophobic interactions between variable loops across protomers at the trimer apex [20].

While escape from PGT145 occurred via eliminating the epitope's positive charges, escape also occurred via *introducing* charges at sites where the wildtype residue is not charged. These included sites 123, 124, and 127 in or very near the epitope, as well as more distant sites encircling the epitope, including sites 200, 202, 203 in the β3-β4 loop, and 312, 313, and 315 at the tip of the V3 loop (Figure 3.6F). These mutations presumably also affected the charge repulsions at the trimer interface and/or overall trimer conformation, disrupting the electrostatic balancing act that is crucial for PGT145 binding.

Escape from CD4bs bnAbs

Escape from CD4bs bnAbs VRC01 and 3BNC117 occurred in both the canonically defined CD4bs epitope and other sites distal to the CD4 binding site (Figure 3.7). In the CD4 binding site, mutations to site 279 in loop D and site 369 in the CD4 binding loop escaped both antibodies (Figure 3.7A). With the exception of sites 279 and 280, the specific amino-acid mutations in loop D that mediated escape differed between VRC01 and 3BNC117.

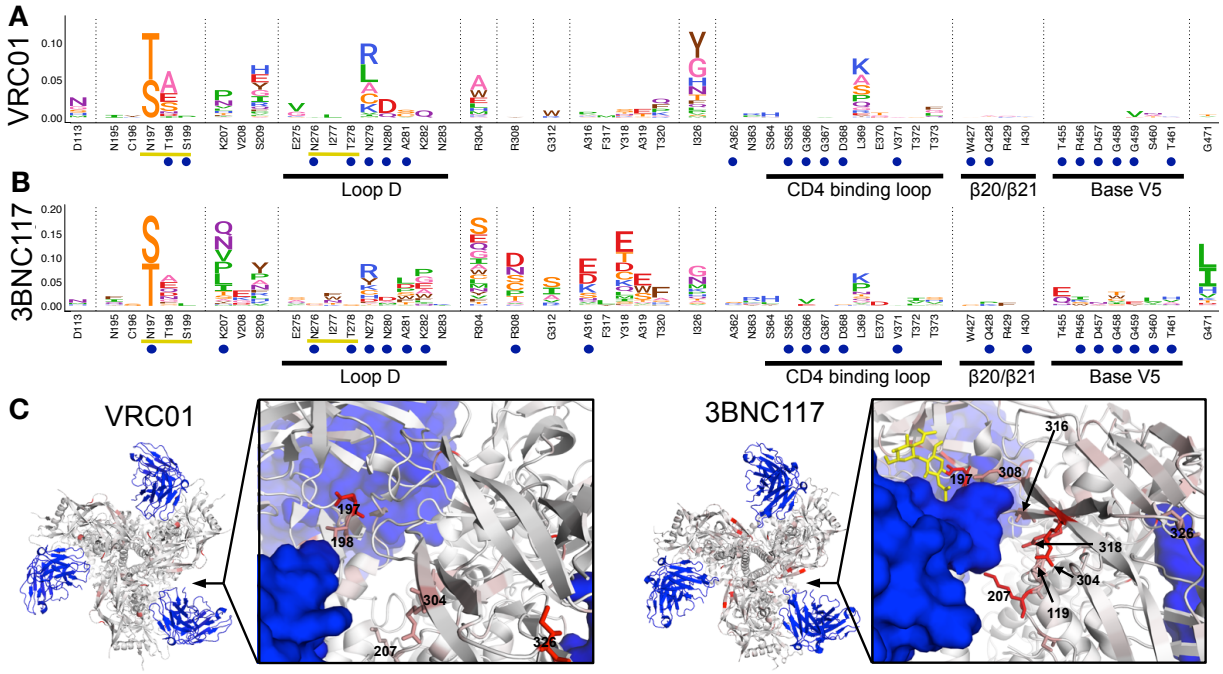


Figure 3.7. Escape from CD4bs bnAbs. A,**B.** Escape profiles for CD4bs bnAbs VRC01 and 3BNC117. Letter heights indicate the excess fraction surviving for each mutation. Blue circles indicate structurally defined contact sites, and yellow underlines indicate a N-linked glycosylation motif. Portions of the canonical CD4bs epitope are underlined in black and labeled. **C.** Antibodies are shown in blue, and Env is colored according to the average fraction surviving at each site (PDBs: 5FYK and 5V8M respectively).

The largest-effect mutations for both antibodies introduced a serine or threonine in place of the asparagine at site 197. The N197 glycan is part of a glycan fence that shields the CD4bs [113]. N197S/T both eliminate this N197 glycan and introduce a new potential N-linked glycosylation motif (PNG) at N195. Since escape was only mediated by S/T at site 197, these data suggest that eliminating N197 alone does not result in viral escape, but shifting the N197 glycan to N195 does. We validated these observations using point mutants in TZM-bl neutralization assays: simply eliminating the N197 PNG via N197E resulted in ~50 fold *more* potent neutralization by VRC01, while N197S resulted in viral escape, increasing the IC_{50} by 27-fold relative to wildtype (Figure 3.5).

Escape from both antibodies also occurred via D113N, which introduces a PNG at site 113. Site 113 is in the trimer interface distant from the CD4bs (Figure 3.7C), suggesting this mutation may affect exposure of the CD4bs epitope by altering trimer conformation or dynamics. We validated that D113N resulted in escape from VRC01 using a TZM-bl neutralization assay (Figure 3.5). While it is unknown if the PNG created by D113N is indeed glycosylated, these data show that altering potential glycosylation sites both near and distal to the epitope can affect CD4bs bnAb neutralization.

Escape from 3BNC117 also occurred at numerous sites near where the antibody's HCDR3 arm makes inter-protomer contacts, including sites in V3 (sites 304, 308, 312, 316-320), and at the base of the β3-β4 loop (sites 207, 209) (Figure 3.7B). This quaternary nature of the 3BNC117 epitope was first postulated based on early trimer structures [114], and higher resolution cryo-EM of BG505 trimer in complex with 3BNC117 [20] confirmed that 3BNC117 directly interacts with residues 207, 308, and 316 from the neighboring protomer (Figure

3.7B,C). It has been previously reported that mutations to site 207 result in decrease 3BNC117 binding [115]. We also observed viral escape from 3BNC117 at site I326, a site distal to 3BNC117 near the base of the V3 loop that takes part in variable loop hydrophobic interactions that may regulate trimer dynamics [20].

Strikingly, while VRC01 does not make similar inter-protomer structural contacts as 3BNC117 [7], we still observed escape at sites 207, 209, 304 and 326 (Figure 3.7A, C). We validated that I326Y results in escape from VRC01 (Figure 3.5), but has little effect on the V3-specific bnAbs 10-1074 and PGT121, despite these antibodies directly contacting this site.

Escape from fusion peptide and gp120/gp41 interface bnAbs

Maps of escape from PGT151 and VRC34.01 highlight the complex nature of the conformational fusion peptide and gp120/gp41 interface epitope (Figure 3.85A,B). Here, we reanalyzed VRC34.01 mutational antigenic profiling data from a prior study [55] quantifying the effects of mutations using the *fraction surviving* metric rather than the *differential selection* metric used in the earlier study, and compared these data to the BG505 Env escape from PGT151 reported here. While both antibodies contact the 6 N-terminal residues of the fusion peptide (512-517), escape from PGT151 is focused on just the 3 N-terminal residues of this peptide (512-514), while escape from VRC34.01 is mediated by numerous mutations to sites 512-516 and 518. The structural footprints of both antibodies center on the fusion peptide, but they contact distinct glycans and protein regions of gp120 and gp41. Again, their *functional epitopes* include distinct subsets of these of protein residues and glycans (Figure 3.8A,B). For both antibodies, there are also numerous sites of significant escape at non-contact sites near the epitope (Figure

3.8A,B). For PGT151, we also identified sites of escape at more distant residues in V3; the mechanisms of escape at these sites are unclear.

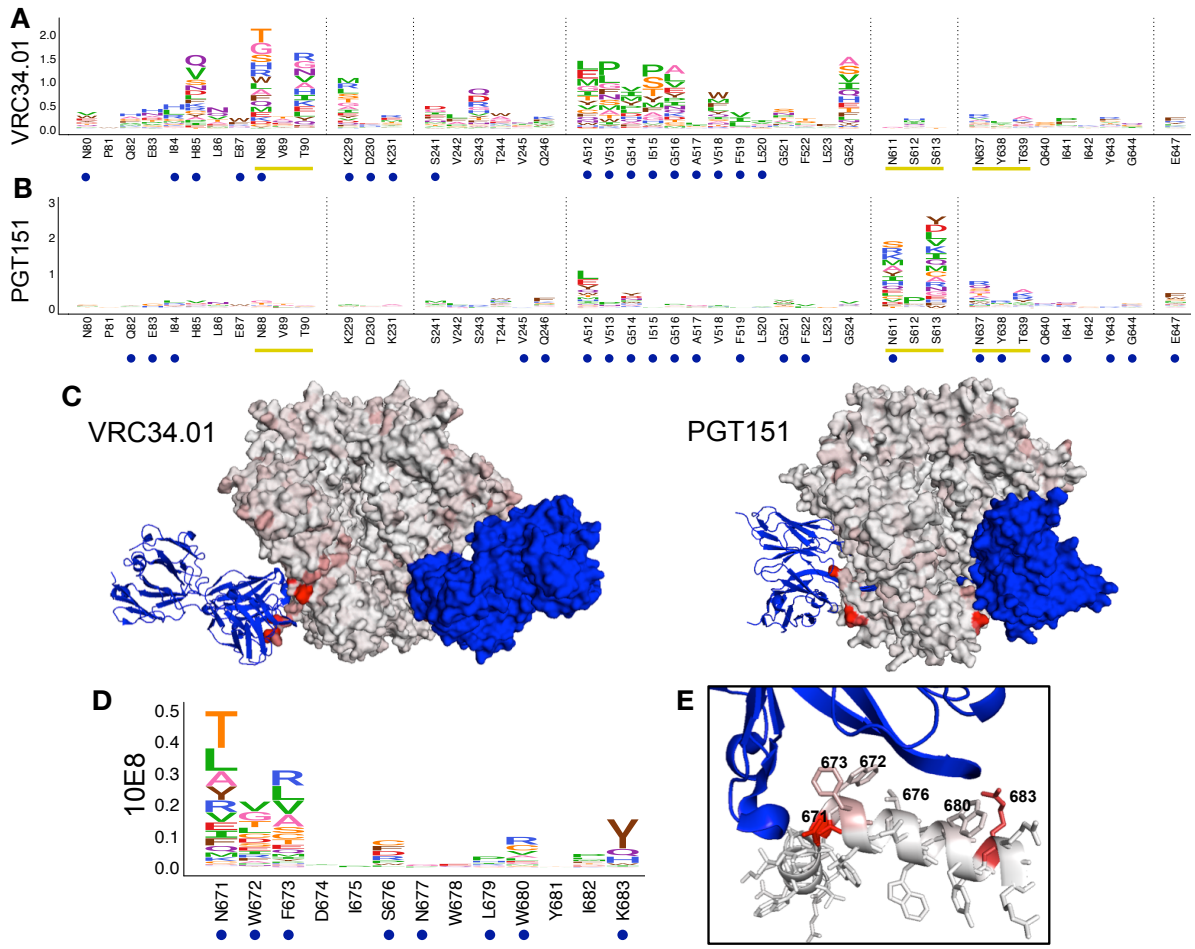


Figure 3.8. Escape from fusion peptide and MPER bnAbs. A, **B.** Escape profiles for fusion peptide and gp120/gp41 interface bnAbs VRC34.01 and PGT151. Letter heights indicate the excess fraction surviving for each mutation. **C.** Fusion peptide antibodies are shown in blue, and Env is colored according to the maximum excess fraction surviving at each site (PDBs: 5I8H and 5FUU respectively). **D.** Escape profile for MPER bnAb 10E8, presented in the same manner as A, B. **E.** 10E8 is shown in blue, and the MPER peptide is colored according to the maximum excess fraction surviving at each site (PDB 4G6F).

Escape from an MPER bnAB

Escape from the MPER-directed antibody 10E8 occurred predominantly in the structurally defined contact sites, with sites of escape localizing to one side of the MPER peptide α-helix apical to 10E8 (Figure 3.8D,E) [103]. This agrees precisely with prior studies [103,116]. However, we identified two additional modest but significant sites of escape outside of the MPER peptide, at sites 609 and 643 in the C-C loop and HR2 domain of gp41 respectively (Figure 3.2). Mutations at these sites may alter fusion kinetics and/or the presentation of the MPER epitope.

Comparing our mutational maps with in vivo *escape during bnAb immunotherapies in humans*

Several of the bnAbs that we characterized have been used in human immunotherapy studies. Some of the escape mutations identified in our work overlap with mutations that arose in the humans during these studies (Table 3.1). For example, when 10-1074 was administered to HIV-infected individuals, viral escape mutations emerged at site 325 and the PNG that encompasses sites 332 and 334 [46]. These are the same three sites where the strongest selection is observed in our 10-1074 mutational antigenic profiling (Table 3.1).

Table 3.1 **Overlap between mutational antigenic profiling sites of escape and sites of viral evolution that occurred *in vivo* during bnAb immunotherapy**

Table 3.1. Overlap between mutational antigenic profiling sites of escape and sites of viral evolution that occurred *in vivo* during bnAb immunotherapy

Antibody	Study	ClinicalTrials.gov Identifier (NCT Number)	Study description	Overlap between **reported** sites of viral evolution during bnAb immunotherapy* and significant sites of viral escape from our mutational antigenic profiling
10-1074	Casket et al 2017	NCT02511990	Single infusion of 10-1074 to HIV infected individuals	**325**, **332**, **334**[A]
3BNC117	Caskey et al 2015, Schoofs et al 2016	NCT02018510	Single infusion of 3BNC117 to HIV infected individuals	**209**, **279**, **308**, **318**, **471**[B]
3BNC117	Scheid et al 2016	NCT02446847	Multiple (2-4) infusions of 3BNC117 after the discontinuation of antiretroviral therapy	**274**[C]
VRC01	Lynch et al 2015a	NCT01950325	One or two infusions of VRC01 in antiretroviral treated and untreated HIV-infected patients, respectively	none[D]
VRC01	Bar et al 2016	NCT02463227 and NCT02471326	Multiple (3-8) infusions of VRC01 after the discontinuation of antiretroviral therapy	**279**[E]**;** reanalysis of these data revealed potential selection at **326****

* The sites in this table were determined in in each study individually, using disparate methods, explained below.

A: Discussed in main figures and text.

B: Used LASSIE (Longitudinal Antigenic Sequences and Sites from Intrahost Evolution) to identify sites selected within the 24-week time frame, using a selection cutoff changing ≥ 80% amino acid frequency compared to baseline. Indels are omitted from this group.

C: The mutations identified here are solely ones discussed in text based on sequence alignments.

D: Used a neutralization-based epitope prediction (NEP) algorithm to predict mutational differences that could be associated with VRC01 selection. The top 5 highest scoring sites for the four patients examined are reported here.

E: Two separate criteria were used by the two clinical trials reported in Bar et al 2016. In one, the VRC01 antibody footprint sequence was analyzed using LASSIE, and we then identified sites within the examined regions that differed ≥50% amino acid frequency compared to baseline. In the other study, HIV env sequences were analyzed using a neutralization-based epitope prediction (NEP) algorithm. Changes in amino-acid residues that occurred within or next to the VRC01 epitope were reported.

**Examination of *env* genotype from this study in the context of our VRC01 escape profile revealed a potential additional site of viral escape no discussed in Bar *et al*. 2016. In Patient V10, T326 was fixed shortly after VRC01 treatment, but T326 was present in only 2/49 sequences at later timepoints when antibody levels may have waned (I326, present in the remaining sequences, is 96.8% conserved in the

LANL filtered web alignment) (Bar et al., 2016). No pretreatment sequences were available.

There is also considerable overlap between sites of escape we map *in vitro* and those that occurred *in vivo* during treatment of infected individuals with the CD4bs antibodies 3BNC117 or VRC01. For 3BNC117, the sites that overlap between our maps and human trials [43,117] include sites 182, 209, 279, 308, 318, and 471 (Table 3.1). Site 279 is one of the strongest sites of escape from 3BNC117 and VRC01 in our experiments. A mutation to site 279 was part of the viral escape pathway within the patient from whom VRC01 was isolated [80], and arose during VRC01 immunotherapy post treatment interruption [44]. Mutations to site 279 also played a role in escaping a CD4bs targeted response in another patient [118] and during 3BNC117 immunotherapy of infected individuals [43,117].

Intriguingly, our data may also be useful for identifying previously unappreciated escape mutations during immunotherapy. For example, after patient V10 underwent therapy with VRC01 [44] a rare amino acid variant at site 326 was fixed in the viral population (Table 3.1), but the potential significance of this mutation was not noted in the original publication since it is far from the structural epitope. Our mutational antigenic profiling shows that mutations at site 326 increase resistance to VRC01, demonstrating how comprehensive maps of mutational escape can aid in interpreting clinical data.

Escape from pooled antibodies

Many immunotherapy studies are beginning to treat patients with combinations of bnAbs. For instance, a recently completed set of clinical trials involved treating patients with equal

concentrations of 3BNC117 and 10-1074 [119,120]. We therefore investigated how escape from a mix of these two antibodies compares to escape from each antibody individually.

We pooled the antibodies at equal concentrations, and then selected our viral libraries with the antibody pool (Figure 3.9, 3.10). Escape from the pooled antibodies appeared to be a combination of the escape profiles from each antibody in isolation (Figure 3.9A, B). For example, we observed escape at sites 325, 332, and 334, likely associated with escape from 10-1074, as well as escape at sites 304, 308, and 471, which presumably affect 3BNC117 resistance.

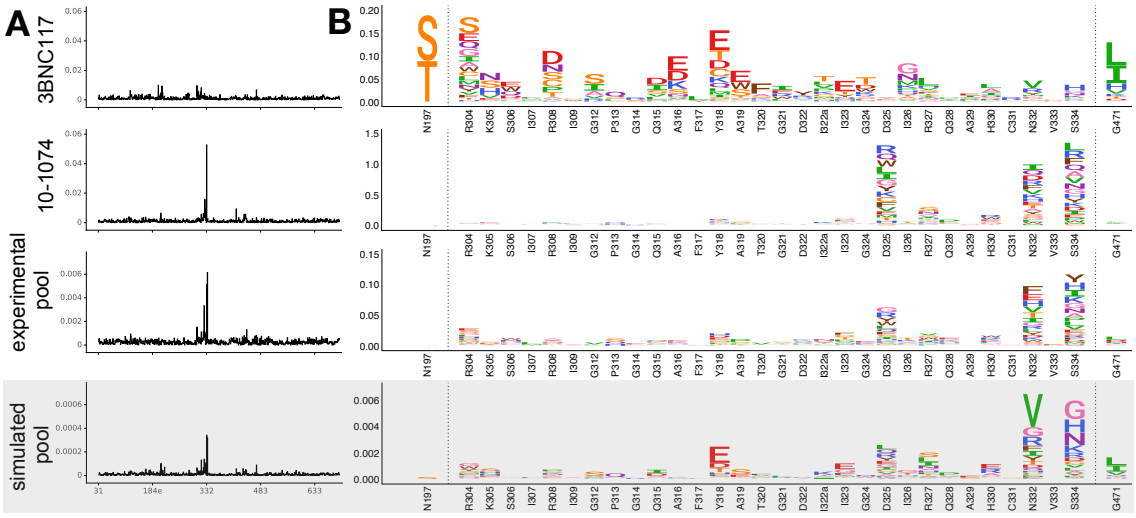
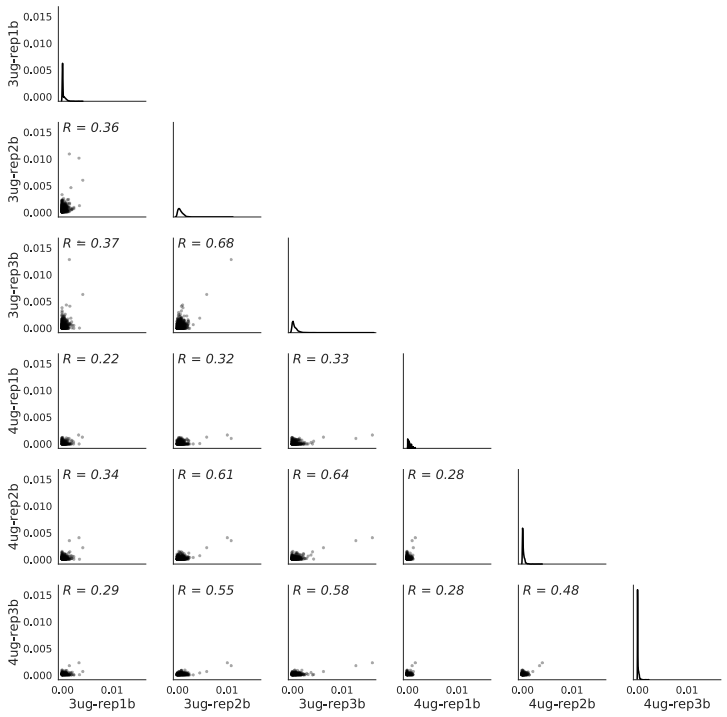


Figure 3.9. Escape from a 3BNC117 and 10-1074 pooled bnAbs. A. The excess fraction surviving neutralization averaged across all mutations at each site. Data from Figure 3 (10-1074) and Figure 4 (3BNC117) are re-plotted for relevant sites. For the pooled 3BNC117 and 10-1074 data, the mean value across six replicates is plotted. The simulated data is the product of each antibody's mean excess mutation fraction surviving values. **B.** A logoplot zooming in on epitope regions for each dataset. In **A** and **B**, the simulated data is distinguished from the experimental data with a light grey overlay.



antibody	antibody concentration (each) per mL	replicate	fraction survivng	percent neutralized
3BNC117 and 10-1074 pooled	3ug	1b	0.00142	99.858
	3ug	3b	0.001521	99.8479
	4ug	1b	0.001074	99.8926
	4ug	3b	0.000413	99.9587
	4ug	2b	0.000341	99.9659
	3ug	2b	0.001887	99.8113

Figure 3.10. The correlation between mutational antigenic profiling biological replicates for the pooled 3BNC117 and 10-1074 antibodies. A. The correlation between the excess fraction surviving averaged across mutations at each site for each biological replicate of the pooled antibody selections. **B.** For each biological replicate, the antibody concentration used during the selection, which mutant virus library was used, and the fraction of that library that survived antibody selection is shown. For clarity, the percent neutralized (1- library fraction surviving) × 100 is also shown.

Importantly, escape from the pooled antibodies does not occur at sites where the escape mutation from one antibody sensitizes the virus to the other. For instance, the strongest escape

mutations for 3BNC117 alone are N197S and N197T, which shift the N197 glycan to N195 (3.9A). However, eliminating the N197 glycan *increases* the virus's susceptibility to neutralization by antibodies targeting the same epitope as 10-1074 [121,122]. Mutations at site N197 are not selected by the pool of antibodies in our mutational antigenic profiling, presumably because any benefit with respect to escaping 3BNC117 is canceled out by increased susceptibility to 10-1074. This example demonstrates the potential for suppressing viral escape mutations by selecting antibodies with synergistic effects at specific sites.

To model the synergistic effects of antibodies on suppressing viral escape, we calculated the expected escape profile from a pool of 10-1074 and 3BNC117 by simply taking the product of each mutation's excess fraction surviving value for each antibody (Figure 3.9). The rationale behind this calculation is that the expected fraction of virions with a mutation that should survive both antibodies is simply the product of the fraction that would survive each antibody individually. The escape profile predicted by this simple model closely matches the actual selection from the pooled antibodies (Figure 3.9).

Overall, these data suggest that no single amino-acid mutations robustly escape both 3BNC117 and 10-1074. Rather, the low level escape from the pooled antibodies appears to represent mutations that escape one antibody but have little effect on the other. Furthermore, the similarity of the experimentally measured escape profile for the pooled antibodies and the profile predicted from the product of the individual antibody profiles suggest that our maps of escape from single antibodies could be useful for computationally predicting the potential for escape from antibody pools.

Quantifying the ability of the BG505 Env to escape each bnAb with single mutations

Anti-HIV bnAbs are often evaluated in terms of their breadth and potency against panels of naturally occurring viral strains. Our data offer the opportunity to calculate an alternative measure relevant to the potential efficacy of bnAb immunotherapies: the ability of single amino-acid mutations to increase the antibody resistance of a particular viral strain.

We used our mutational antigenic profiling to assess the ease of single-mutation escape of the BG505 Env from each antibody. First, we simply qualitatively examined the 100 largest effect-size mutations for each antibody (Figure 3.11A). For some antibodies (such as VRC34.01), there are many individual mutations that efficiently escape neutralization—but for other antibodies (such as VRC01), only a few mutations affect escape, and do so with only moderate size effects (Figure 3.11A). Interestingly, for all of the antibodies, at least some of the largest effect-size mutations are accessible by single nucleotide mutations, indicating that the genetic code only has moderate affects on the accessibility of escape mutations (Figure 3.11A).

To quantify the ease with which the BG505 Env can escape each antibody by single mutations, we summed the excess mutation fraction surviving values at each significant site of viral escape. This single-mutation ease of escape metric is moderately correlated with antibody's breadth (Figure 3.11B). But especially for the broadest antibodies, there are differences between neutralization breadth on natural strains and ease of single-amino acid escape by BG505 (Figure 3.11B). For example, 10E8 has 98% breadth and VRC01 has 91% breadth, but BG505 has more capacity to escape 10E8 by single amino-acid mutations. Differences in the extent of natural sequence variation at epitopes potentially contributes to the differences between breadth and the ease of escape: many of the escape mutations we identify are rarely observed in nature [123]. These results highlight that, similar to influenza antibodies [106], HIV bnAb breadth on natural strains and the potential for single-mutation escape by any given viral variant are distinct

measures, both of which may be useful for assessing the potential for viral antibody escape in clinical settings.

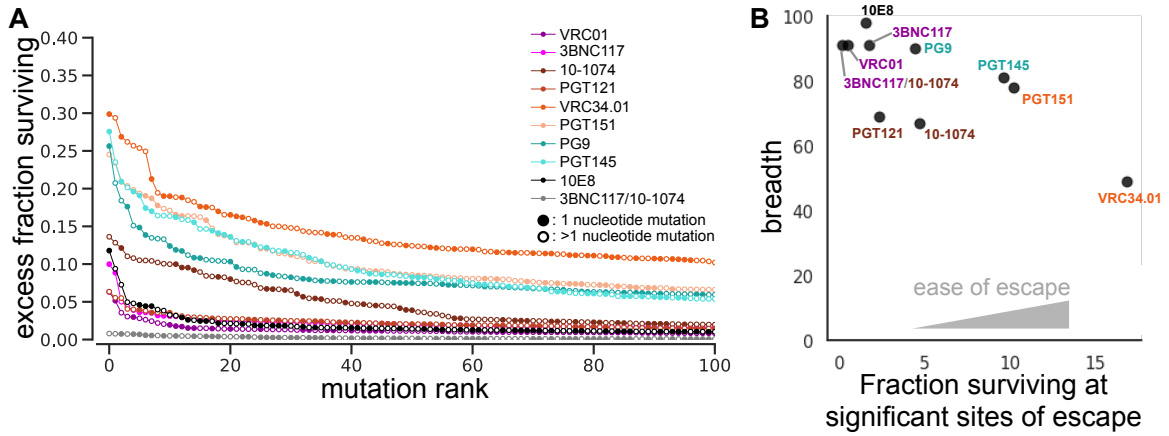
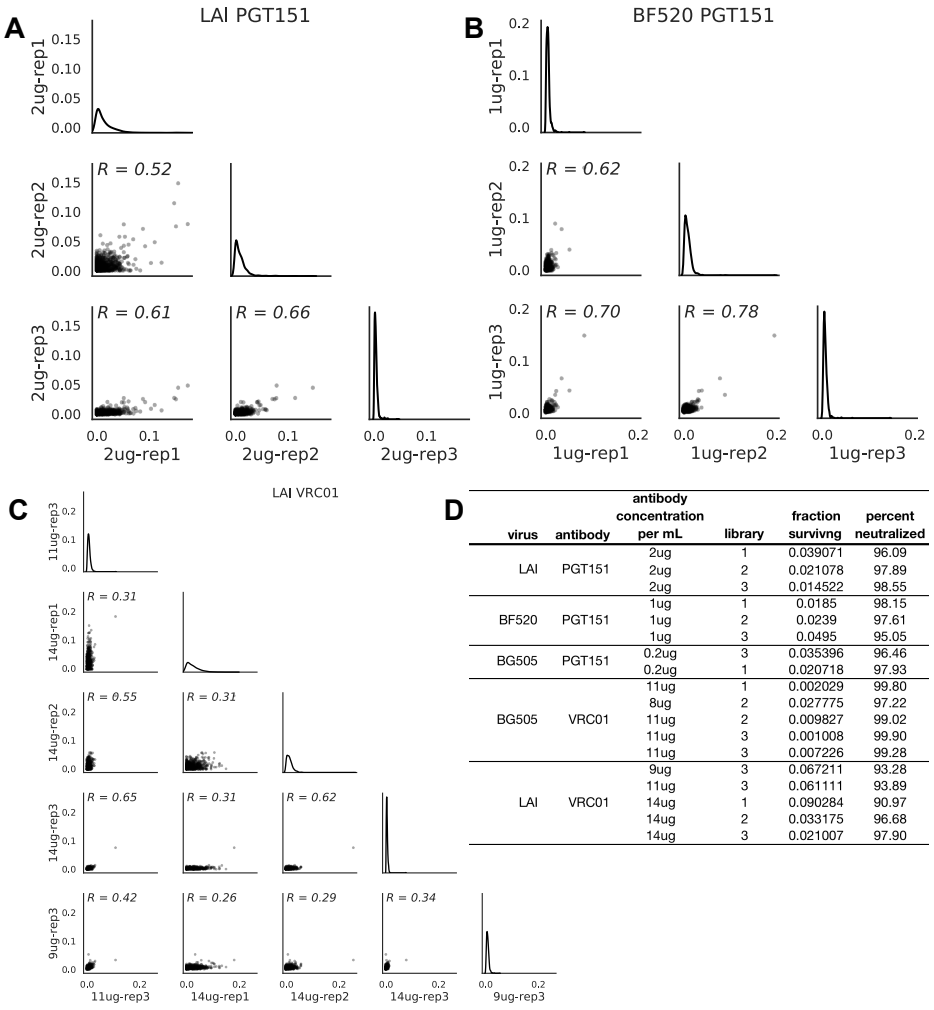


Figure 3.11. Quantifying the ability to escape each bnAb. A. The excess fraction surviving for the 100 largest effect mutations for each antibody. Closed circles indicate mutations that are one nucleotide mutation away from BG505 wildtype, while open circles indicate mutations that are 2 or 3 nucleotide mutations away from BG505 wildtype sequence. **B.** Each antibody's breadth (as quantified in Figure 2) is plotted against the sum of the excess mutation fraction surviving values at all significant sites of viral escape.

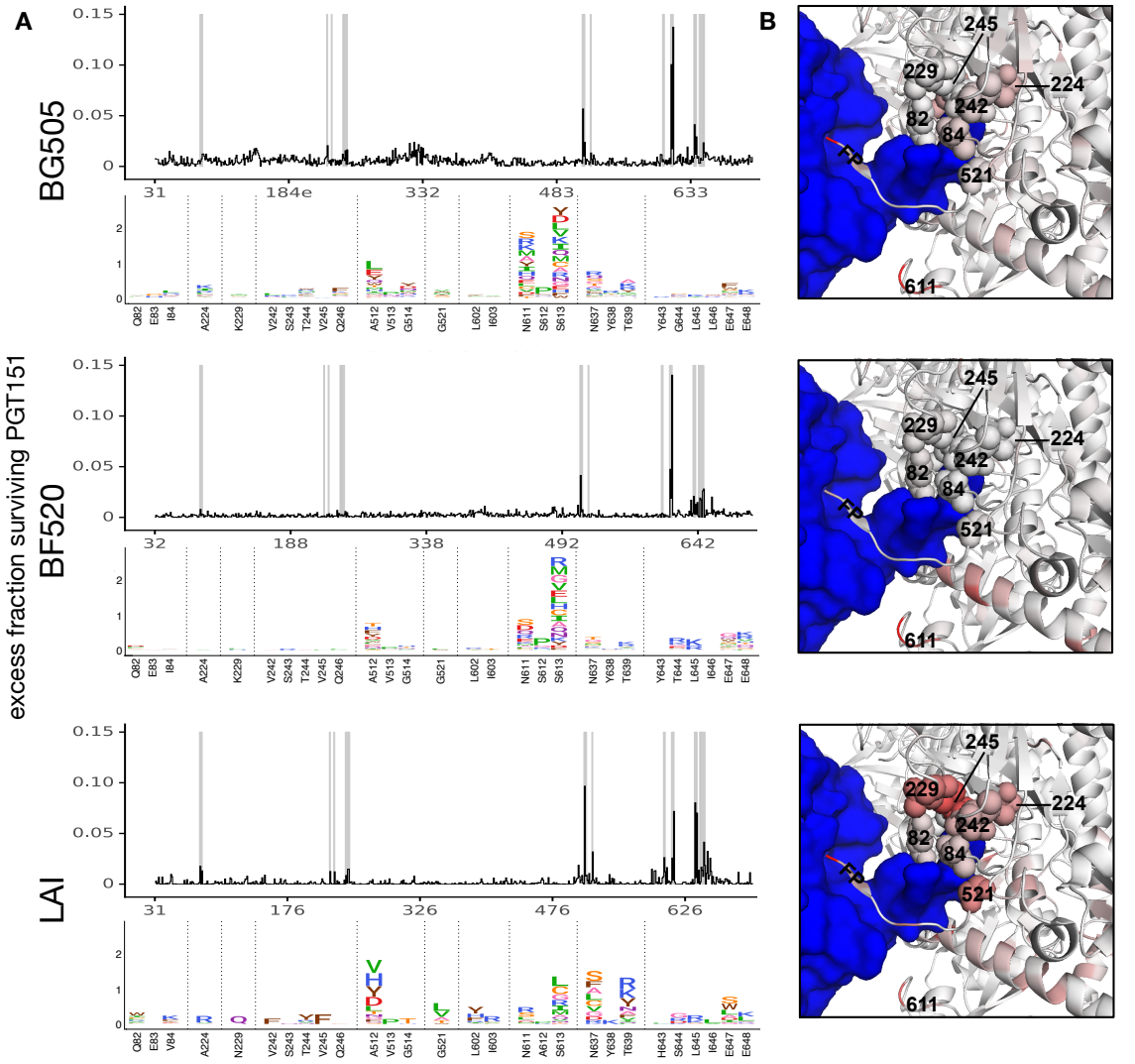
Mapping differences in escape across diverse Envs

All of the foregoing results are for Env from a single viral strain, BG505. While it is well known that the effect of mutations on antibody neutralization can differ from strain to strain, there are no unbiased quantifications of how much the antigenic effects of mutations typically shift across strains. To address this question, we also mapped escape from VRC01 and PGT151 using LAI Env, a subtype B strain (Figure 3.12-3.14). In addition, we compared out maps of escape from PGT151 in BG505 and LAI to our previously published map of escape from this antibody by the subtype A BF520 Env [53].



virus	antibody	antibody concentration per mL	library	fraction survivng	percent neutralized
LAI	PGT151	2ug	1	0.039071	96.09
		2ug	2	0.021078	97.89
		2ug	3	0.014522	98.55
BF520	PGT151	1ug	1	0.0185	98.15
		1ug	2	0.0239	97.61
		1ug	3	0.0495	95.05
BG505	PGT151	0.2ug	3	0.035396	96.46
		0.2ug	1	0.020718	97.93
BG505	VRC01	11ug	1	0.002029	99.80
		8ug	2	0.027775	97.22
		11ug	2	0.009827	99.02
		11ug	3	0.001008	99.90
		11ug	3	0.007226	99.28
LAI	VRC01	9ug	3	0.067211	93.28
		11ug	3	0.061111	93.89
		14ug	1	0.090284	90.97
		14ug	2	0.033175	96.68
		14ug	3	0.021007	97.90

Figure 3.12. The correlation between mutational antigenic profiling biological replicates for PGT151 and VRC01 using LAI and BF520 mutant Env libraries.. A. The correlation of the excess fraction surviving averaged across mutations at each site between biological replicates for PGT151 escape using LAI mutant Env libraries. **B.** Same as A, but for BF520 mutant Env libraries. **C.** The same as A, but for escape from VRC01 using LAI mutant virus libraries. **D.** For each biological replicate, the antibody concentration used during the selection, which mutant virus library was used, and the fraction of that library that survived antibody selection is shown. For clarity, the percent neutralized (1- library fraction surviving) × 100 is also shown.

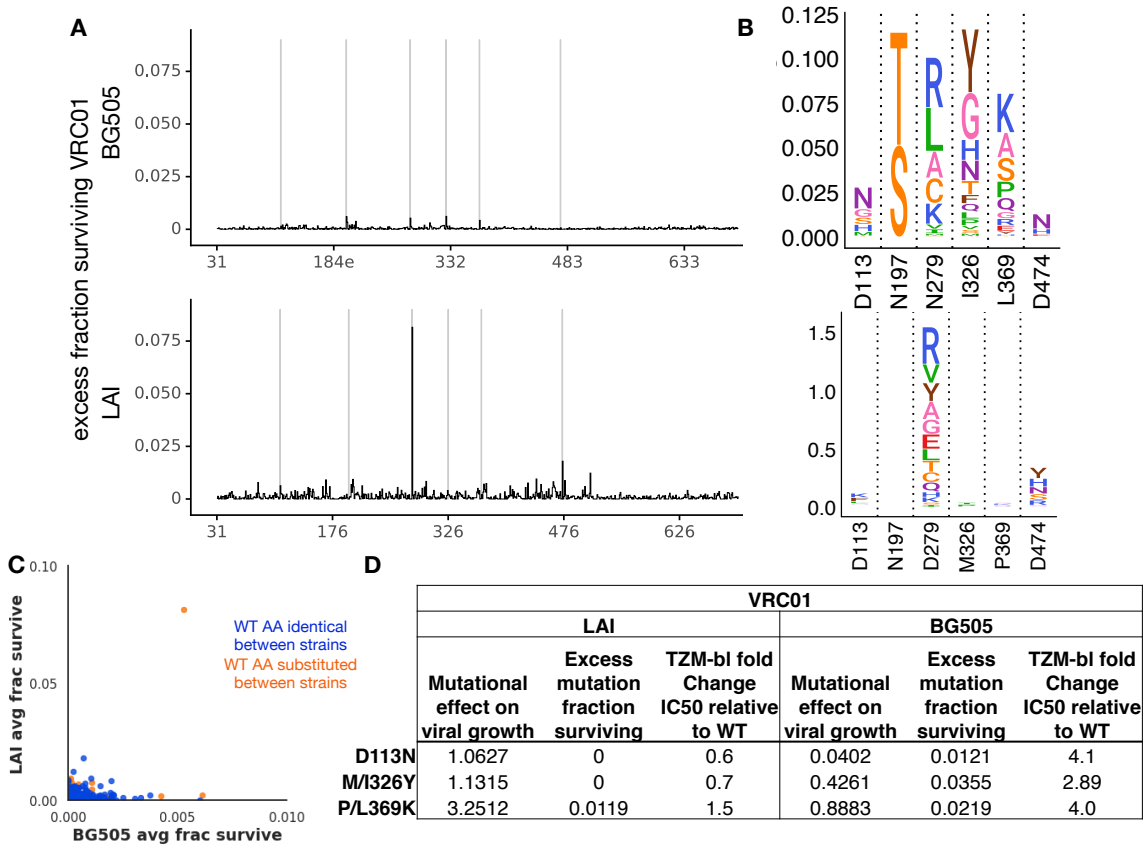


D

	PGT151								
	LAI			BG505			BF520		
	Mutational effect on viral growth	Excess mutation fraction surviving	TZM-bl fold Change IC50 relative to WT	Mutational effect on viral growth	Excess mutation fraction surviving	TZM-bl fold Change IC50 relative to WT	Mutational effect on viral growth	Excess mutation fraction surviving	TZM-bl fold Change IC50 relative to WT
V/I84K	1.9583	0.1153	4.6	0.2852	0.0138	2.1	0.1942	0.0073	0.8
G521L	0.0350	0.2216	>12	0.0075	0	>200	0.0029	0.0351	>7.25
I603R	0.3261	0.1539	>14.3	0.0481	0.0099	2.7	0.0176	0	4.0

Figure 3.13. Differences in PGT151 escape across Envs. A. The line plots shows each site's average excess fraction surviving PGT151 for BG505, BF520, and LAI Env. Beneath, logoplots

show the mutation level escape profile for sites highlighted in grey in the line plots. **B.** View of the PGT151 epitope colored according to escape in each Env. PGT151 Fab is colored blue, and Env is colored according to the maximum mutation fraction surviving at each site for each strain (PDB:5FUU). **C.** The correlation between BG505, BF520, and LAI excess fraction surviving PGT151 for each site. Sites are colored blue if the wildtype amino acid is identical between strains, orange if they differ. **D.** The mutational effect, excess mutation fraction surviving, and fold change in PGT151 IC_{50} from TZM-bl assays using point mutant pseudoviruses of validated point mutants. As in Figure S4, the mutational effect is measured via deep mutational scanning of each Env under selection for viral replication in cell culture in the absence of any immune selection (Haddox et al 2016, Haddox et al 2018). The mutational effect is the ratio of the preference of that mutant relative to the wildtype amino acid preference at that site. If the mutational effect is >1, then that mutation replicates better than wildtype in cell culture; if it is <1, that mutant replicates worse than wildtype. Where there are discrepancies between mutational antigenic profiling and TZM-bl assays, that mutation has an effect <<1, indicating it is deleterious for viral replication in cell culture. For example, TZM-bl neutralization data validate that V/I84K has different antigenic effects on PGT151 across strains. In contrast, between-strain PGT151 differences in the excess fraction surviving at sites 521 and 603 may be shaped by differences in mutational tolerance. In mutational antigenic profiling, I603R mediated escape in LAI but not in BG505 or BF520; similarly, G521L had differing effects across strains. However, these pseudovirus point mutants escaped PGT151 in all three strains, albeit to differing extents. Because we assayed only replication-competent mutant viruses, these mutations were likely purged from our BG505 and BF520 libraries during functional selection in cell-culture before antibody selection, while they remained in our LAI libraries.



	VRC01					
	LAI			BG505		
	Mutational effect on viral growth	Excess mutation fraction surviving	TZM-bl fold Change IC50 relative to WT	Mutational effect on viral growth	Excess mutation fraction surviving	TZM-bl fold Change IC50 relative to WT
D113N	1.0627	0	0.6	0.0402	0.0121	4.1
M/I326Y	1.1315	0	0.7	0.4261	0.0355	2.89
P/L369K	3.2512	0.0119	1.5	0.8883	0.0219	4.0

Figure 3.14. Differences in VRC01 escape across Envs. A. The line plots shows each site's average excess fraction surviving VRC01 for BG505 and LAI Env. **B.** Logoplots show the mutation level escape profile for sites highlighted in grey in A. **C.** Same as in S8C, but for VRC01 escape in BG505 and LAI. **D.** Same as in S8D, but for validated VRC01 escape mutants. The differences in VRC01 escape across Envs may be explained by differences in conformation or dynamics. BG505 is generally well-behaved as a SOSIP trimer, adopting a prefusion "closed" state [124], and LAI is a lab adapted strain which likely adopts a more "open" conformation similar to the related strain NL4-3 [9]. We speculate the mutations outside of the CD4bs, such as introducing a PNG at the trimer interface (D113N) or altering hydrophobic interactions in variable loops at the trimer apex (I326Y) could disrupt BG505's trimer structure or dynamics

and result in escape from VRC01, but the same mutations to LAI may not disrupt a "closed" state if it is less frequently or not adopted by LAI.

Escape from PGT151 across all three strains is quite similar. The dominant escape mutations are in the N611 and N637 PNGs, as well as sites in the fusion peptide and HR-2 domain (Figure 3.13). However, there are a number of sites where single mutations mediated escape in LAI but not in BG505 or BF520 (sites 82, 84, 224, 229, 242, 245, 521, 602, 603). Many of the escape mutations unique to LAI are clustered near the epitope, suggesting that LAI may adopt a different local conformation (Figure 3.13B). We validated three of these strain-specific escape mutations using TZM-bl neutralization assays and found that the strain differences in both antigenic effects and mutational tolerance appear to play a role in shaping these differences (Figure 3.13D).

There were relatively more strain-specific differences in escape from VRC01 than from PGT151 (Figure 3.14). Mutations to site 279 escaped VRC01 in both Envs, but with larger effects in LAI (Figure 3.14B). N197S/T escaped VRC01 in BG505, but not in LAI. There is a clear mechanistic basis for this difference: N197S/T shifts the N197 PNG to N195 in BG505, while N197S/T only disrupts the N197 PNG in LAI. In contrast, D113N introduces a glycan in both strains, but only escaped VRC01 in BG505 (Figure 3.14B, validated in TZM-bl assays in Figure 3.14D). Further, I326Y and L369K escaped VRC01 in BG505 but not LAI (also validated in TZM-bl assays in Figure 3.14D), while mutations to site 474 generally have a larger effect in LAI than in BG505 (Figure 3.14B). A possible explanation for the differences in escape is that these Envs may adopt different conformations with different dynamics [9].

Discussion

We have mapped how all single amino-acid mutations to the BG505 Env affect neutralization of replication-competent HIV by nine prototypical bnAbs. These maps of viral escape define the functional epitopes of these antibodies, which we show are distinct from the structurally defined epitopes. For all antibodies, viral escape occurred at only a fraction of structurally defined contact sites, and many escape mutations occurred at sites outside the direct structurally defined epitope. This escape at non-contact sites often clustered close to the structural contacts, suggesting that altering a network of interacting sites near the structural epitope can disrupt antibody binding or neutralization. In a few cases, escape also occurred at sites more distant from the epitope, likely due to alterations in the conformation or dynamics of Env.

Others have previously noted that not all structurally defined contact sites affect antibody binding [59,60,97], and mutations outside of the structural epitope can affect antibody sensitivity [125–130]. However, our complete maps of escape mutations make it possible to systematically quantify the overlap between the functional and structural epitopes of bnAbs. Our work also highlights how incompletely prior structural and virological assays have defined the functionally relevant interactions between HIV and antibodies. Even though our study focused on some of the best-characterized anti-HIV bnAbs, we uncovered numerous sites in Env (often outside of the structural epitope) where mutations were not previously known to affect bnAb sensitivity.

One application for which our maps of viral escape may be of immediate use is evaluating bnAb immunotherapies. During immunotherapy trials, many viral mutations arise, and it is important to know which ones alter sensitivity to the bnAb used in the trial [43–47,117].

While our maps of course do not perfectly predict escape that occurs *in vivo*, which could be influenced by many stochastic factors including the specific variant that infected the individual, they do define functional epitopes that can be used to help assess the potential antigenic significance of viral mutations. For instance, our map of the functional epitope of VRC01 allowed us to re-interpret the potential significance of a previously unremarked upon mutation outside the structural epitope that occurred during VRC01 immunotherapy.

We also examined how our maps of escape from individual antibodies compared to those generated using a pool of antibodies—an important question since combinations of bnAbs are a clear future direction in antibody immunotherapy [131,132]. We mapped escape from a pool of two bnAbs, 3BNC117 and 10-1074, and we found that there were no mutations that robustly escaped both antibodies. This finding agrees with the results of two recently completed clinical trials that administered this antibody combination and did not report selection of any mutations that conferred resistance to both antibodies [119,120]. Further, the combined selection of the two antibodies was similar to that predicted from the product of the individual escape profiles. Therefore, antigenic maps for individual antibodies may be useful for modeling viral escape from combinations of antibodies.

Our data also provide a way to quantify the ease with which the BG505 Env can escape from each antibody via single mutations. While this ease of escape metric contains many caveats specific to our experimental system (it considers only single mutations to BG505 that support viral replication in cell culture, omitting the effect of insertions and deletions), caveats also apply to other methods for estimating how likely a virus is to escape an antibody. For example, characterizing viral escape during immunotherapy in animal models often examines only a single viral strain in a limited number of animals. We found that BG505's ease of escape by single

mutations is correlated with antibody breadth against natural sequences, but that there are also differences, especially among the broadest antibodies. We suggest that both measures may be useful for assessing the potential efficacy of bnAb-based therapies.

Of course, it is important to keep in mind that our maps only measure the effect single amino-acid mutations to BG505. Epistatic interactions among multiple mutations can play a role in viral fitness and immune evasion [80,133–138]. Similarly, we performed most of our experiments in the single genetic background of the BG505 Env, but the effects of mutations on viral growth and antigenicity sometimes differ among viral strains [52,74,139]. However, we did compare maps of escape from a few antibodies across viral strains. Escape from PGT151 was quite similar across strains, while escape from VRC01 was more variable between strains. Further studies contrasting escape from additional antibodies across additional strains are needed to better define strain-specific effects in antibody escape.

Nonetheless, having complete antigenic maps of the BG505 Env versus a panel of important bnAbs provides a wealth of information that can help guide the study of HIV evolution and the development of anti-viral strategies. Future work that combines these antigenic maps with measurements [51,52] or models [140–142] of how mutations affect HIV's replicative fitness could shed further light on the virus's evolutionary dynamics under immune pressure.

Methods

Generation of mutant virus libraries

We have previously described the BG505.T332N mutant proviral DNA libraries and the resulting functional mutant virus libraries [52]. Briefly, triplicate mutant BG505.W6M.C2.T332N *env* libraries that contained randomized codon-level mutations to sites

31-702 (HXB2 numbering is used here and throughout this manuscript) were independently generated and cloned into Q23.BsmBI.ΔEnv proviral plasmid [52]. These proviral plasmid libraries, as well as wildtype proviral plasmid, were transfected into 293T cells (obtained from ATCC). We then passaged the transfection supernatant at an MOI of 0.01 TZM-bl infectious units/cell in SupT1.CCR5 cells. The resulting genotype-phenotype linked mutant virus libraries were concentrated via ultracentrifugation.

Mutational antigenic profiling

The Env mutational antigenic profiling approach has been previously described [53,55]. Briefly, 0.5 - 1×10^6 TZM-bl infectious units of independent mutant virus libraries were neutralized with each antibody at an ~IC_{95} - $IC_{99.9}$ concentration for one hour. Libraries were then infected into 1×10^6 SupT1.CCR5 cells in R10 (RPMI, supplemented with 10% FBS, 1% 200 mM L-glutamine, and 1% of a solution of 10,000 units/mL penicillin and 10,000 μg/mL streptomycin) containing 100ug/mL DEAE-dextran. In parallel to each antibody selection, each mutant virus library was also infected into 1×10^6 SupT1.CCR5 cells without antibody selection to serve as the experiment-specific non-selected control. Four 10-fold serial dilutions of each mutant virus library were also infected into 1×10^6 cells as an infectivity standard curve. Cells were resuspended in 1 mL fresh R10 at three hours post infection. Cells were washed with PBS, pelleted, and non-integrated viral DNA was isolated via a miniprep at twelve hours post infection. The amount of viral genome in each sample was quantified via qPCR [89] or ddPCR [53], and the fraction of each selected library that survived antibody selection relative to the non-selected control was interpolated from the infectivity standard curve.

Barcoded subamplicon deep sequencing

The *env* gene was then amplified and sequenced using a barcoded subamplicon sequencing approach as previously described [52] and explained in more detail at https://jbloomlab.github.io/dms_tools2/bcsubamp.html. Briefly, we first amplified the entire *env* gene from the harvested non-integrated viral DNA. This full-length *env* amplicon was then used as a template for amplifying 7 subamplicons that tile across *env.* Each of these subamplicons contained primer-introduced random barcodes on each end (8×N). Subamplicons are then bottlenecked such that the number of unique ssDNA molecules is less than the sequencing depth and then subjected to a final round of PCR that added the remainder of the Illumina sequencing adapters. All amplicons were then pooled and sequenced on Illumina HiSeq 2×250 bp runs. Errors introduced during sequencing were corrected by taking the consensus sequence at each site for each uniquely tagged ssDNA molecule.

Structural analyses

Antibody contact sites were defined from Env-antibody structural models. The PDB models used were: 5FYK for VRC01 [7], 5V8M for 3BNC117 [20], 5T3Z for 10-1074 [143], 3U4E for PG9 [111], 5V8L for PGT145 [20], 5FUU for PGT151 [73], 5I8H for VRC34.01 [25], 4G6F and for 10E8 [103]. High resolution models of PGT121 bound to Env are not available; we used a model of PGT122 (PDB: 5FYL) [7], which is a closely related "PGT121-like" clonal variant of PGT121 [101]. Contact residues were defined as any Env residue where a non-hydrogen atom comes within 4 Å of any non-hydrogen antibody atom. When an antibody contacted a glycan, the N of that glycan's PNG was counted as a contact. For asymmetric

antibody-trimer structures, the closest distance of the three antibody-Env residue distances was used.

We used the above referenced models (with one exception) to generate figures using PyMol. While we used the high resolution PG9-V2 scaffold structure [111] to determine contact sites, we mapped the fraction surviving values onto the moderate resolution model of PG9 in complex with BG505 trimer (PDB: 5VJ6) [112] in Figure 3F to better illustrate the quaternary aspect of this apex epitope.

TZM-bl neutralization assays

TZM-bl neutralization assays using BG505.T332N pseudoviruses bearing single additional point mutants were performed and analyzed as previously described [53]. The assay was performed in duplicate two or three independent times, and fold change in IC_{50} of each mutant relative to BG505.T332N wildtype pseudovirus was calculated independently for each experiment and then averaged across all replicates.

Analysis of deep sequencing data and computation of fraction surviving.

We used dms_tools2 version 2.2.9 (https://jbloomlab.github.io/dms_tools2/) to analyze the deep sequencing data and calculate the fraction surviving [94]. The calculation of the fraction surviving statistic has been described previously [106] and is documented in detail at https://jbloomlab.github.io/dms_tools2/fracsurvive.html. Sequencing of wildtype proviral DNA plasmid was used as the error control during the calculation of the fraction surviving. We took the median values across all experimental replicates for each antibody, and plotted the excess

fraction surviving data on logoplots rendered by dms_tools2 using weblogo [93] and ggseqlogo [144].

Identification of significant sites of viral escape

Since the signal to noise ratio appeared to differ between antibodies, we defined statistically significant sites of viral escape beyond background individually for each antibody. For each antibody, we fit a gamma distribution to binned site fraction surviving values (median values across all replicates) using robust regression (soft L1 loss as implemented in scipy). We then identified sites that fell outside the range of values expected from this distribution at a false discovery rate of 0.01. For the purposes of Figures 2B and 2C, multiple sites that disrupted a PNG that the antibody contacted was considered a single site of escape. However, when the antibody also contacted the second or third protein residues in the PNG and that site was site of significant escape, the site treated as additional site of escape.

Data and software availability

Open-source software to analyze mutational antigenic profiling datasets is available at https://jbloomlab.github.io/dms_tools2/. The computational analysis is provided at https://github.com/jbloomlab/EnvsAntigenicAtlas. Illumina deep sequencing reads are available from the NCBI SRA as study SRP157948, BioProject PRJNA486029, accession numbers SRR7693968- SRR7694021, SRR7758666, and SRR8168127-SRR8168140.

Chapter IV

Complete functional mapping of infection- and vaccine-elicited antibodies against the fusion peptide of HIV

The text in this chapter has been modified slightly from: PloS Pathogens 2018 Jul 5; 14 (7), e1007159. doi: 10.1371/journal.ppat.1007159

Summary

Eliciting broadly neutralizing antibodies (bnAbs) targeting envelope (Env) is a major goal of HIV vaccine development, but cross-clade breadth from immunization has only sporadically been observed. Recently, Xu et al (2018) elicited cross-reactive neutralizing antibody responses in a variety of animal models using immunogens based on the epitope of bnAb VRC34.01. The VRC34.01 antibody, which was elicited by natural human infection, targets the N terminus of the Env fusion peptide, a critical component of the virus entry machinery. Here we precisely characterize the functional epitopes of VRC34.01 and two vaccine-elicited murine antibodies by mapping all single amino-acid mutations to the BG505 Env that affect viral neutralization. While escape from VRC34.01 occurred via mutations in both fusion peptide and distal interacting sites of the Env trimer, escape from the vaccine-elicited antibodies was mediated predominantly by mutations in the fusion peptide. This functional data, in combination with structural data, suggest that the breadth of vaccine-elicited antibodies targeting the fusion peptide can be enhanced by specific interactions with additional portions of Env. Thus, our complete maps of viral escape both delineate pathways of resistance to these fusion peptide-directed antibodies and provide a strategy to improve the breadth or potency of future vaccine-induced antibodies against Env's fusion peptide.

Introduction

The isolation of broadly neutralizing antibodies (bnAbs) capable of neutralizing diverse clades of HIV-1 has invigorated hopes of developing a broadly protective antibody-based HIV vaccine [145,146]. Epitope mapping of bnAbs has revealed a number of conserved sites of vulnerability on Env, and designing immunogens to elicit antibodies that target such sites is a promising vaccine strategy [18,147]. Structural characterization of the epitope of bnAb N123-VRC34.01 (subsequently referred to as VRC34.01) revealed that the N terminus of the fusion peptide (FP) is one such site of vulnerability [25]. A number of additional bnAbs that also partially target this epitope have been characterized [72,74–76,148], suggesting it may be a promising vaccine target.

Xu et al (2018) [26] used the VRC34.01 epitope as a template to design vaccine approaches that elicited antibodies to the fusion peptide capable of neutralizing diverse strains of HIV-1. The broadest of these vaccine-elicited fusion peptide-directed antibodies (vFP antibodies) derive from a class of murine antibodies, the vFP1 class, whose members share similar B cell ontogenies and structural modes of recognition. While the murine vFP1 class antibodies share VH gene sequence homology with the human VH gene of VRC34.01, the binding mode and antibody contact sites of the vFP antibodies are different from that of VRC34.01 [26]. Here we focus primarily on two of these antibodies, 2712-vFP16.02 and 2716-vFP20.01, which have breadths of 31% and 27% respectively, on a 208-strain cross-strain panel [26]. The vFP16.02 antibody was elicited by a BG505 Env trimer prime and three fusion peptide-carrier protein conjugate (FP-KLH) boosts in two-week intervals, and vFP20.01 was elicited via a similar scheme with an additional trimer boost. Antibodies were isolated via the generation of hybridomas three weeks after the last boost. The breadth of these antibodies rivals

infection-elicited bnAbs such as 2G12 [149], although these vFP antibodies have lower affinity for Env than the original VRC34.01 bnAb [26]. To improve breadth and potency in further rounds of structure-based vaccine design, we undertook studies to define precisely the epitope specificities of the vFP antibodies, to delineate viral pathways of functional resistance, and to understand how these differ from the original template VRC34.01 bnAb.

Results

We used mutational antigenic profiling [53] to quantify how all single amino-acid mutations at each residue of Env affected the neutralization of replication-competent HIV by VRC34.01 and the two vaccine-induced antibodies, vFP16.02 and vFP20.01. Specifically, for each antibody, we neutralized libraries [52] of viruses carrying all amino-acid mutations to the ectodomain and transmembrane domain of the BG505.T332N Env at an ~IC_{95} antibody concentration. For each of the possible 12,730 Env mutations (670 mutagenized sites × 19 amino acid mutations), we calculated the enrichment of the mutation in the antibody-selected condition relative to a non-selected mutant virus library, a quantity that we term *differential selection* [53]. This entire process was performed in full biological replicate, beginning with independent generation of the proviral plasmid mutant libraries. The results from the two replicates were well correlated. For the remainder of the paper, we present the averaged data from these two replicates.

It was immediately apparent that viruses with mutations in the N terminus of the fusion peptide (sites 512-516) robustly escaped all three antibodies (Figure 4.1A). However, there were also obvious differences between VRC34.01 and the two vaccine-elicited antibodies. While viral escape was concentrated in the fusion peptide for the vFP antibodies, there were a number of

additional sites in Env where mutations allowed for escape from VRC34.01. Prior structural and functional characterization of VRC34.01 have revealed that it interacts with both the N terminus of the fusion peptide and the N-linked glycan at residue 88 [25]. The results of these prior structural and functional studies are highly correlated with our mutational antigenic profiling data. Mutational antigenic profiling also revealed additional sites of viral escape from VRC34.01, including residues 84, 85, 227, 229, 241 and 243, which are all proximal to the fusion peptide (Figure 4.1B). We also observed a modest but reproducible signal for escape from vFP16.02 and vFP20.01 via introducing a glycosylation motif at N241, where a glycan would cause steric hindrance to the approach of these antibodies, and via mutations at site 85 (Figure 4.2A).

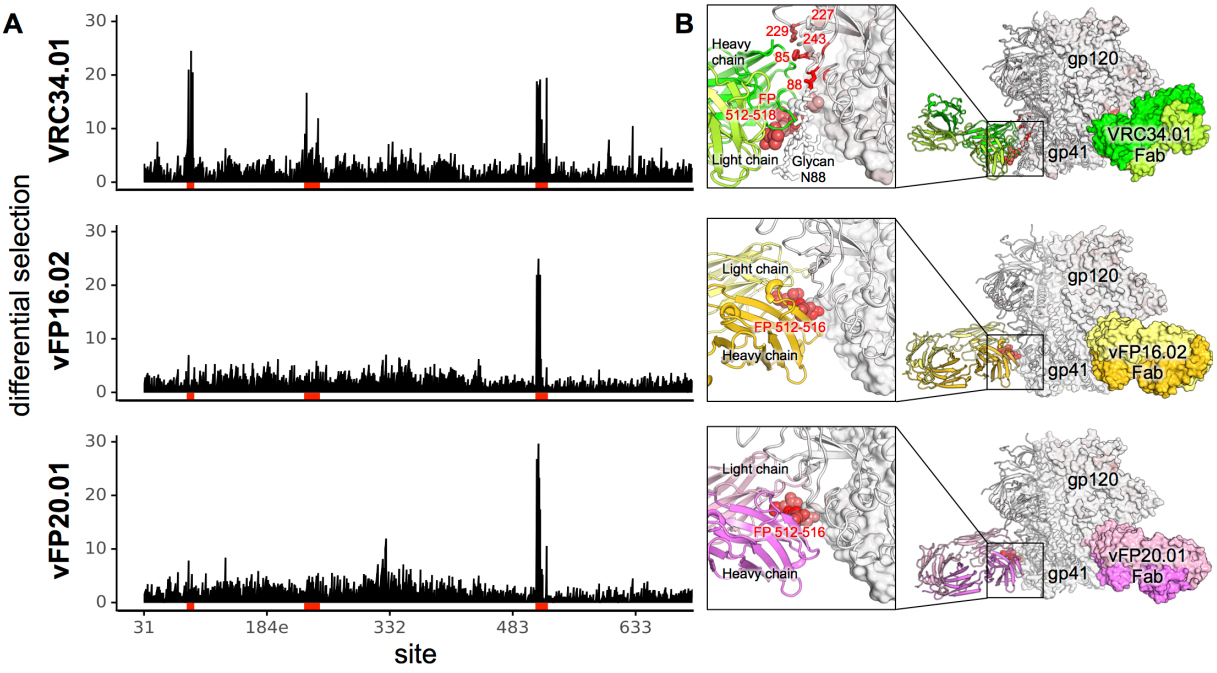


Figure 4.1. Complete escape profile of natural and vaccine elicited fusion peptide antibodies. A. The positive site differential selection is plotted across the length of the mutagenized portion of Env for each antibody. **B.** Structural representation of the FP-directed antibody epitope on BG505 Env trimer. The Env trimer is colored from white to red according to

the positive differential selection at each site. Cartoon diagram of Fab-trimer complexes were adapted from crystal structure (VRC34.01) and cryoEM structure (vFP16.02 and vFP 20.01), with one Fab-protomer shown in ribbon and the other two Fab-protomers shown as surface. Labeled fusion peptide residues are shown in spheres. Detail of the Fab-trimer interaction was shown in the inset.

Since we measured the effect of every amino acid at every site, we could examine viral escape at a *mutation* rather than *site* level (Figure 4.2A). For example, V518W, V518M, and V518L resulted in drastic escape from VRC34.01 neutralization, while V518A had no effect [25]. Further, while the mutational antigenic profiling results agreed with alanine and glycine scans of the fusion peptide [26], the complete escape profiles identified additional escape mutants at nearly all of the fusion peptide residues that contact each antibody. This mutation-level mapping was especially informative for understanding the mechanistic basis for escape from neutralization. For example, hydrophobic amino acids were among the most enriched escape mutants at site 512-516 for all three antibodies. This may reflect their ability to eliminate antibody binding while still allowing for the fusion peptide to be efficiently inserted into the host cell plasma membrane during fusion.

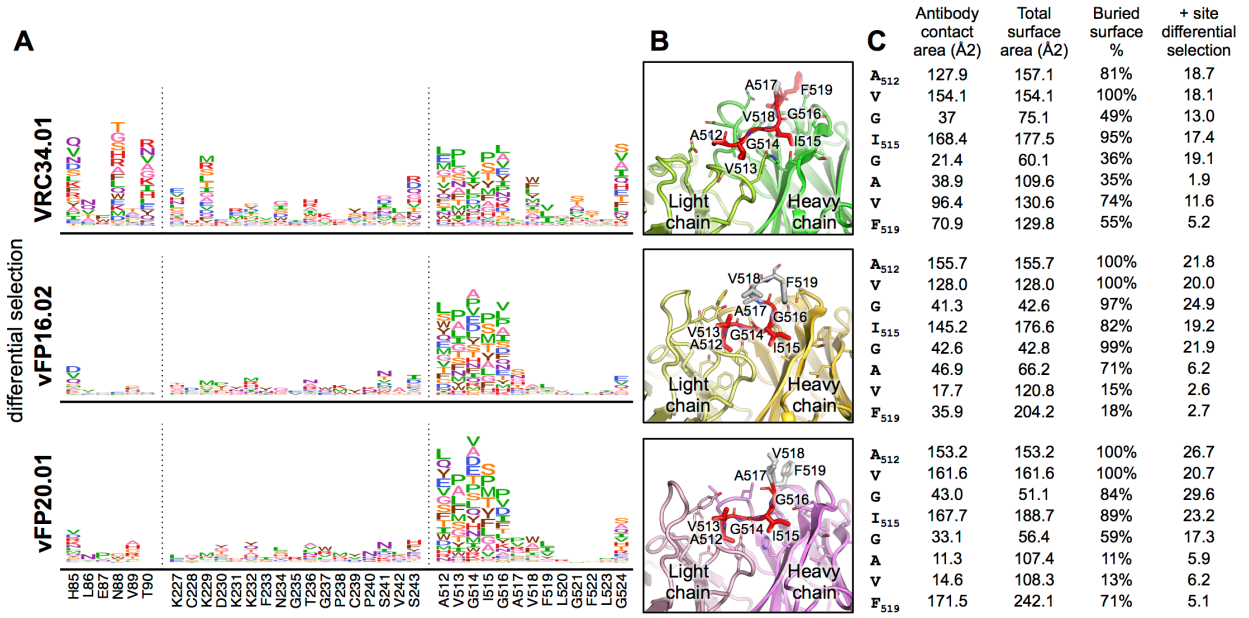


Figure 4.2. The mutation-level escape profiles agree with structural characterizations of the fusion peptide epitope. A. The mutation-level escape profile for regions with considerable differential selection by at least one antibody (regions with red underlay in Fig 1A). Left to right, these include the N88 glycosylation motif and neighboring area, surface exposed gp120 residues that interact with VRC34.01, and most of the fusion peptide. The height of each amino acid is proportional to the logarithm of the relative enrichment of that mutation in the antibody selected condition relative to the non-selected control. **B.** Fab-FP peptide crystal structures of each antibody. Antibody Fabs were colored as in Fig. 1B, FP residues with significant differential selection to each antibody were colored in red. **C.** Total surface, antibody-contact surface (determined from co-crystal structures shown in Fig 2B), and positive site differential selection for each antibody

Even within the fusion peptide region, there are differences between the template, infection-elicited bnAb and the vaccine-elicited antibodies. Escape from VRC34.01 could also occur via mutations at sites 518 and 521, whereas these mutations had little impact on the

vaccine-elicited antibodies. These residue level differences in escape within the FP agree with structural interactions; within FP peptide-Fab co-crystal structures, the percent buried surface area of each residue tracks well with the extent of escape for each antibody (Figure 4.2B, C). Likewise, the fusion peptide dominated the epitope of antibodies vFP16.02 and vFP20.01 bound to Env, with the cryo-EM reconstructions showing strongest electron density of the bound antibody around the fusion peptide [26], with residues 512-519 accounting for ~63% and ~56% of the total protein epitope surface for vFP16.02 and vFP20.01, respectively. However, these data also highlighted how not all structural contacts contribute equally to binding. Cryo-EM (vFP16.02 and vFP20.01) and crystal (VRC34.01) structures of each Fab bound to HIV-1 Env showed contacts of the antibodies with residues 517, 519, and 520 of the fusion peptide of the trimeric Env (Figure 4.1B). Mutational antigenic profiling suggested that these residues were less critical to escape than residues 512-516.

We also found an unexpected site of viral escape at site 524 that was shared between all three antibodies. Although Gly 524 is distal to the epitope and does not directly contact any of the antibodies, it is part of a structural motif conserved in all three antibodies, involving a tight turn that allows Phe 522 to anchor the flexible fusion peptide via van der Waals interactions to α6 helix of gp41, and β5 and β7 strands of gp120 [25,26]. Interestingly, mutations at site 524 mediated escape from each antibody to differing extents (Figure 4.2A). Thus, these distal mutations likely altered the presentation or conformational dynamics of the N terminus of the fusion peptide, an observation relevant to immunogen design.

Mutational antigenic profiling also identified mutations distal to the epitope that were depleted, rather than enriched, during antibody selection relative to the non-selected libraries (Figure 4.3). These include mutations that eliminated the N611 glycosylation motif, indicating

viruses that lack the N611 glycan are more sensitive to neutralization by these antibodies. To a lesser extent, we observed the same phenotype at the N88 glycosylation motif for the vFP antibodies but not for VRC34.01 (Figure 4.3A,B). These results are consistent with the fact that binding of VRC34.01 is dependent on the N88 glycan [25], and disrupting the N88 and N611 glycosylation motifs individually in BG505 pseudoviruses results in more potent neutralization by vFP16.02 and vFP20.01 [26]. Numerous mutations throughout Env were depleted during VRC34.01 selection, while far fewer mutations were depleted during vFP16.02 and vFP20.01 selection (Figure 4.3A). For all three antibodies, sites 621 and 624 were amongst those under the strongest negative differential selection (Figure 4.3A,B). These surface-exposed gp41 sites, which neighbor the FP epitope but do not directly interact with any of the antibodies, could allosterically affect the conformation and/or presentation of the fusion peptide or neighboring glycans. Indeed, site 621 is part of a surface salt bridge network that could be important for maintaining local structure [26]. While the mechanism of this phenomenon is yet to be fully elucidated, introducing these mutations into Env trimers may be a way to increase the exposure of the fusion peptide in trimer immunogens.

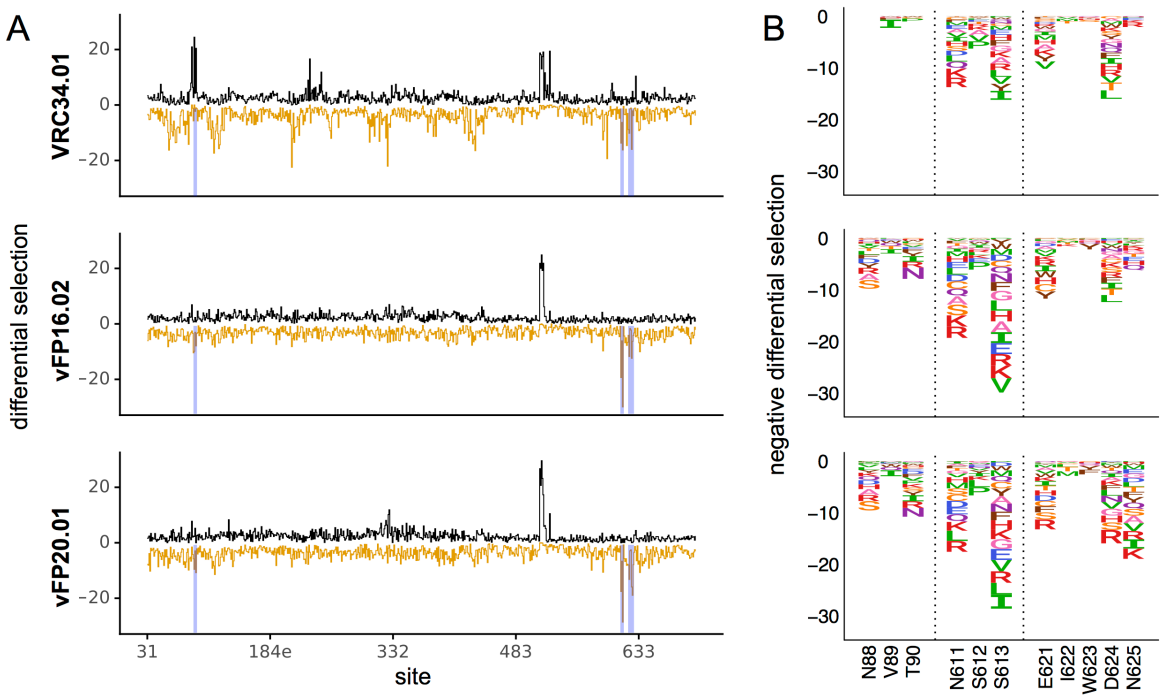


Figure 4.3. Mutations that better expose the fusion peptide are selected against during antibody treatment. A. The positive (black) and negative (orange) site differential selection is plotted across the length of the mutagenized portion of Env for each antibody. **B.** The negative differential selection for regions of interest (regions highlighted in blue in 1A). Left to right, these include the N88 glycosylation motif, the N611 glycosylation, and surface exposed gp41 sites that consistently have mutations depleted upon antibody selection. The height of each amino acid is proportional to the logarithm of the relative depletion of that mutation in the antibody selected condition relative to the non-selected control.

Our mutational antigenic profiling may also help to contextualize prior structural analyses [25,26]. vFP antibodies have a restricted angle of approach and substantial interactions with glycans at residues N88, N295, N448, and N611 [26], but these data show that eliminating these glycans did not result in viral escape. Thus, these antibodies likely approach the fusion

peptide at an angle that accommodates or avoids steric clashes with glycans. We hypothesize that the vFP antibodies' lack of direct reliance on glycans is a result of the vaccination regimen, where responses were induced with FP-KLH immunogens that lacked appropriately positioned glycans.

Further, vFP16.02 and vFP20.01's focused recognition of the fusion peptide's N-terminus is also reflected by their Env-bound cryo-EM reconstructions [26,55]. Reduced electron density for the antibody outside of this region around the fusion peptide indicated greater conformational variability as a result of being less constrained by interactions with Env.

Discussion

Overall, these in-depth delineations of the functional epitopes revealed similarities and differences between how VRC34.01 and the vaccine-elicited antibodies recognize Env. Escape from both vFP16.02 and vFP20.01 was mediated predominantly via mutations in the fusion peptide, whereas VRC34.01 neutralization was also affected by mutations at numerous additional sites.

The focused functional interface of the vaccine-elicited antibodies takes on added significance in the context of a unique structural observation: Env-bound vFP16.02 and vFP20.01 have reduced electron density outside of the paratope contacting the fusion peptide's N-terminus, indicating greater conformational heterogeneity. To further investigate this phenomenon, we determined two more vFP1-class antibody cryo-EM structures. These antibodies also displayed internal disorder, and the extent of relative light chain order correlated with breadth [55]. Increased light chain disorder may indicate weaker interactions with Env gp120. Since our mutational antigenic profiling clearly shows that VRC34.01 but not vFP16.02

or vFP20.01 has functional interactions outside the FP, we hypothesize that additional interactions with other regions of Env may increase the breadth of vFP antibodies. Consistent with this idea, the breadth of vFP antibodies is correlated with their affinity to Env trimer [26]. Our functional mapping does show modest but reproducible selection for escape mutations from vFP16.02 and vFP20.01 at site 85 in gp120 (Figure 4.2A), which stacks against their CDR L3 loops. Thus, it does seem possible for the light chain of vaccine-elicited antibodies to derive advantageous interactions from this region of gp120.

Therefore, mapping *all* mutations that affect neutralization by infection and vaccine-elicited antibodies may help identify potentially desirable interactions to target with trimer boosts or rationale immunogen design. However, such additional interactions outside of the fusion peptide could increase affinity at the cost of reduced breadth if the additional contacts occur at Env sites that are variable in nature. Fortunately, most of the VRC34.01 escape mutations that that are outside of the fusion peptide are at sites largely conserved in nature, suggesting that this tradeoff can be minimized. One strategy to elicit FP antibodies with additional Env interactions is to include additional Env trimer boosts. This could include Δ611 and/or Δ88 glycan trimers, which appear to increase the exposure of this epitope to vFP16.02 and vFP20.01, followed by wildtype trimers in hopes of guiding antibody lineages towards interacting with or avoiding clashes with these glycans. Alternatively, Env trimer immunogens that incorporate other mutations that appear to better expose this epitope (Figure 4.3) could be developed.

An alternative or parallel approach to increasing the breadth of fusion-peptide vaccine responses is to tailor the focused recognition of the fusion peptide to accommodate greater sequence diversity within this region [26]. This goal might be accomplished by immunizing with peptide immunogen pools that incorporate naturally found diversity within these sites [26]. Our

data suggest that a peptide immunogen cocktail designed to elicit antibodies similar to vFP16.02 and vFP20.01 would need only to cover diversity at sites 512-516. Further, rather than designing cocktails that cover *all naturally occurring diversity* at sites 512-516, our data allows for the design of peptide cocktails that cover only those naturally occurring amino acid polymorphisms *that mediate escape from vFP antibodies*.

Within the fusion peptide, hydrophobic amino acids are among the most enriched escape mutations for all three antibodies (Figure 4.2A). This makes sense, since viral fusion requires insertion of Env's hydrophobic fusion peptide into the target cell membrane. However, the numerous escape mutations mapped in our experiments contrast with the high sequence conservation of the fusion peptide's N terminus in nature. The much higher conservation of the fusion peptide in nature may be because there is relatively little natural pressure for mutations at these sites or might indicate the presence of additional constraints in nature that are not present in our cell-culture experiments. Prior work has shown that the BG505 Env used in this study tolerates many mutations at conserved fusion peptide sites under selection for viral replication in cell culture [52]. However, this mutational tolerance varies among strains, since Env from strain BF520 (another subtype A transmitted/founder Env) is much less tolerant of mutations at fusion peptide sides such as 512 and 516 [52].

Translating the promising potential of the these vFP antibodies into consistent, broadly protective responses will require iterative rounds of vaccine design and detailed evaluation of elicited responses. This study is an early step in this process and is limited to examining the neutralization sensitivities and structures of a small number of antibodies. Here, we do not evaluate numerous other antibody-mediated immune functions. Further, it is yet to be determined

if additional Env interactions outside of the FP, similar to VRC34.01, will indeed result in greater breadth, or if increased affinity alone may drive breadth.

In conclusion, our mutational antigenic profiling of vFP antibodies and the template VRC34.01 bnAb, combined with analyses of additional vFP antibody structures, have provided insights that can help refine vaccination regimens. We precisely mapped previously unappreciated interactions between Env and the template bnAb VRC34.01, and our functional and structural analyses suggest that these interactions may play a role in VRC34.01's relatively greater breadth and could possibly help improve the breadth or potency of anti-FP antibodies. Therefore, these complete maps of viral escape provide a detailed atlas to guide subsequent rounds of structure-based vaccine design.

Methods

Antibody production

Mammalian codon-optimized genes encoding either antibody heavy chain or light chain were synthesized at Gene Universal Inc (Newark, DE). Antibody genes were cloned into mammalian expression vector pVRC8400. Antibodies were produced as described in Xu et al (2018) [26].

Generation of mutant virus libraries

The generation and characterization of the BG505 mutant proviral DNA libraries, and the resulting functional mutant virus libraries have been described in detail previously [51,52]. Briefly, proviral DNA libraries that contained randomized codon-level mutations at sites 31-702 (HXB2 numbering) of BG505.W6M.C2.T332N *env* were independently generated. These, and

accompanying wildtype controls, were independently transfected into 293T cells (ATCC). The transfection supernatant was passaged at a low MOI (0.01 TZM-bl infectious units/cell) in SupT1.CCR5 cells (James Hoxie), and the resulting genotype-phenotype linked mutant virus library was concentrated via ultracentrifugation. Env residues are numbered according to HXB2 reference numbering throughout.

Mutational antigenic profiling

We have previously described Env mutational antigenic profiling using BF520 Env libraries [53]; a similar approach was taken here. Briefly, 5×10^5 infectious units of two independent mutant virus libraries were neutralized with VRC34.01, vFP16.02, or vFP20.01 at an ~IC_{95} concentration (33 ug/mL, 500 ug/mL, or 500ug/mL of antibody, respectively) for one hour. Neutralized libraries were then infected into 1×10^6 SupT1.CCR5 cells in R10 (RPMI [GE Healthcare Life Sciences; SH30255.01], supplemented with 10% FBS, 1% 200 mM L-glutamine, and 1% of a solution of 10,000 units/mL penicillin and 10,000 μg/mL streptomycin), in the presence of 100ug/mL DEAE-dextran. Three hours post infection, cells were spun down and resuspended in 1 mL fresh R10 without DEAE-dextran. At 12 hours post infection, cells were washed with PBS and non-integrated viral cDNA was isolated via miniprep. In parallel to each antibody selection, each mutant virus library was also subjected to a mock selection, and four 10-fold serial dilutions of each mutant virus library were also used to infect 1×10^6 cells to serve as an infectivity standard curve. Selected and mock-selected viral cDNA was then sequenced with a barcoded subamplicon sequencing approach as previously described [52]. This approach involves introducing unique molecular identifiers during the Illumina library prep in order to further reduce the sequencing error rate.

The relative amount of viral genome that successfully entered cells was measured using *pol* qPCR[89]. A qPCR standard curve was generated based on each mutant virus library's infectivity standard curve, from which the % neutralization for each antibody selected sample was calculated.

Analysis of deep sequencing data

We used dms_tools2 version 2.2.4 (https://jbloomlab.github.io/dms_tools2/) to analyze the deep sequencing data and calculate the differential selection [94]. The differential selection statistic has been described in detail [53,61] and is further documented at https://jbloomlab.github.io/dms_tools2/diffsel.html. Sequencing of wildtype proviral DNA plasmid was used as the error control during the calculation of differential selection. Differential selection was visualized on logoplots rendered by dms_tools2 using weblogo [93] and ggseqlogo [144].

Data availability and source code

Open-source software that recapitulates the mutational antigenic profiling analysis, starting with the deep sequencing data through the calculation of selection metrics, is available at https://jbloomlab.github.io/dms_tools2/. The entire computational analysis, starting with downloading the deep sequencing reads through generating figures, is available at https://github.com/jbloomlab/MAP_Vaccine_FP_Abs. Illumina deep sequencing reads were deposited into the NCBI SRA as SRR6429862-SRR6429875.

Chapter V

Massively parallel profiling of HIV-1 resistance to the fusion inhibitor enfuvirtide

The text in this chapter has been modified slightly from: bioRxiv 2018 Nov 19; http://dx.doi.org/10.1101/472746

Summary

Resistance to enfuvirtide, the only clinically approved HIV-1 entry inhibitor, has primarily been mapped to the binding site in the N-terminal heptad repeat (NHR) of the Env transmembrane domain and a limited number of allosteric sites. To better delineate the genotypic determinants of resistance, we used deep mutational scanning to quantify how all mutations to HIV-1 Env affect enfuvirtide sensitivity. We identified numerous additional resistance mutations in the NHR and other regions of Env, including the co-receptor binding site. This complete map of resistance sheds light on the diverse mechanisms of enfuvirtide resistance and can inform clinical monitoring of patients.

Introduction

Antiretroviral drug therapy has reduced the global burden of HIV/AIDS. However, HIV-1's exceptional evolutionary capacity enables the virus to evolve resistance, eroding the therapeutic efficacy of drug regimens. Identifying drug resistance mutations is therefore imperative to the clinical management of HIV-1/AIDS, particularly as uptake of antiviral treatment is increasing globally.

Enfuvirtide (also known as T-20 and Fuzeon), the only clinically approved fusion inhibitor, is a 36 amino-acid peptide derived from the C-terminal Heptad repeat (CHR) domain of gp41 that inhibits viral entry into cells. After Env binds CD4, enfuvirtide binds to the coiled-coil NHR structure of the pre-hairpin intermediate and inhibits fusion by preventing the CHR from binding to the NHR to form the 6-helical bundle [6]. Due to enfuvirtide's high cost and need for twice-daily subcutaneous injections, it is used as part of a "salvage therapy" for highly treatment-experienced patients with multidrug resistance. Such patients need to be carefully monitored for resistance before and during therapy [150], which requires knowing the viral genotypic determinants of enfuvirtide resistance.

Prior studies that analyzed enfuvirtide resistance in cell culture [151] and *in vivo* during therapy [152,153] identified resistance mutations at sites 547-556, implicating this region of the NHR as enfuvirtide's binding site. However, mutations in gp120 outside of this region have also been shown to affect enfuvirtide sensitivity, possibly by altering co-receptor tropism, co-receptor affinity, or fusion kinetics [154,155].

Here we more completely map resistance mutations by using deep mutational scanning to quantify how enfuvirtide resistance is affected by all mutations to HIV-1 Env compatible with viral replication.

Methods

Generation of Env mutant virus libraries

We used Env mutant virus libraries from a subtype A transmitted variant BG505 to map enfuvirtide resistance. The generation of these libraries has been previously described [51,52]. Briefly, we independently generated triplicate proviral DNA libraries encoding codon-level mutations to sites 31-702 (HXB2 numbering) of BG505.W6M.C2.T332N *env*. These libraries

contain (670 mutagenized sites) × (19 amino acid mutations) = 12,730 possible amino acid mutations. We then produced full-length viruses from these mutant DNA libraries and passaged them in SupT1.CCR5 cells to select for functional viruses, resulting in mutant virus libraries that encode all functionally tolerated mutations to Env [52].

Resistance profiling

To identify resistance mutations, we incubated the mutant virus libraries with or without enfuvirtide, infected cells, and then identified the mutant viruses that were enriched upon drug selection using deep sequencing. This approach is similar to the mutational antigenic profiling process we have previously used to map antibody escape [53,54]. Briefly, 5×10^5 to 1×10^6 infectious units of three independent mutant virus libraries were incubated in the presence of 8ug/mL of enfuvirtide, then infected into 1×10^6 SupT1.CCR5 cells in R10 (RPMI with 10% FBS, 1% 200 mM L-glutamine, and 100 units/mL of penicillin and streptomycin), containing 100ug/mL DEAE-dextran. Three hours post infection, cells were pelleted and resuspended in 1 mL R10 without DEAE-dextran, and at 12 hours post infection, non-integrated viral cDNA was isolated using a miniprep. As mock-selected controls, each mutant virus library was infected into cells without enfuvirtide selection, as were four 10-fold serial dilutions of each mutant virus library. Selected and mock-selected viral cDNA was then sequenced with a barcoded subamplicon sequencing approach as previously described [52], which introduces unique molecular identifiers used to correct sequencing errors. The amount of virus library that entered cells was quantified via qPCR [54].

Analysis of deep sequencing data

We used dms_tools2 version 2.3.0 (https://jbloomlab.github.io/dms_tools2/) to analyze the deep sequencing data [94]. The differential selection statistic has been previously described [61] and is documented at https://jbloomlab.github.io/dms_tools2/diffsel.html. Sequencing of wildtype proviral DNA plasmid was used as the error control during the calculation of the differential selection.

Data availability and source code

The entire computational analysis is available at https://github.com/jbloomlab/EnfuvirtideResistance. Illumina deep sequencing reads were deposited into the NCBI SRA as SRR8097918- SRR8097920.

TZM-bl inhibition assays

Individual mutations were introduced into BG505.T332N Env, and pseuodviruses were generated with these Envs and tested in TZM-bl neutralization assays using previously described methods.

Results

To quantify the effect of all mutations to Env on enfuvirtide sensitivity, we selected triplicate BG505 mutant virus libraries with 8 ug/mL enfuvirtide. This highly selective drug concentration, which resulted in just 0.15% to 0.78% of the mutant virus libraries surviving selection, was used to mimic the environment in which resistance could arise during therapy. We deep sequenced the *env* genes of viruses that entered cells and calculated the enrichment of each

mutation in the drug-selected condition relative to a non-selected mutant virus library, to calculate a quantity that we term the *differential selection* [61].

Selection of resistance mutations was highly reproducible across biological triplicates. Most resistance mutations were found in the NHR domain (Figure 5.1) and included many well-characterized mutations on the IAS–USA drug resistance mutations list [156]. For example, S553T, a known resistance mutation, was the largest effect size mutation at site 553 in our experiments. Similarly, V549E and V549A were the largest effect size mutations at site 549 in our experiments, and both have been previously shown to cause resistance. However, V549M was not enriched in our experiments despite being associated with resistance in other strains. Similarly, the known I548V resistance mutation was not enriched in our experiments, but numerous other mutations at site 548 were strongly enriched. These discrepancies may be due to strain specific differences in enfuvirtide sensitivity, discussed in more detail in the Discussion.

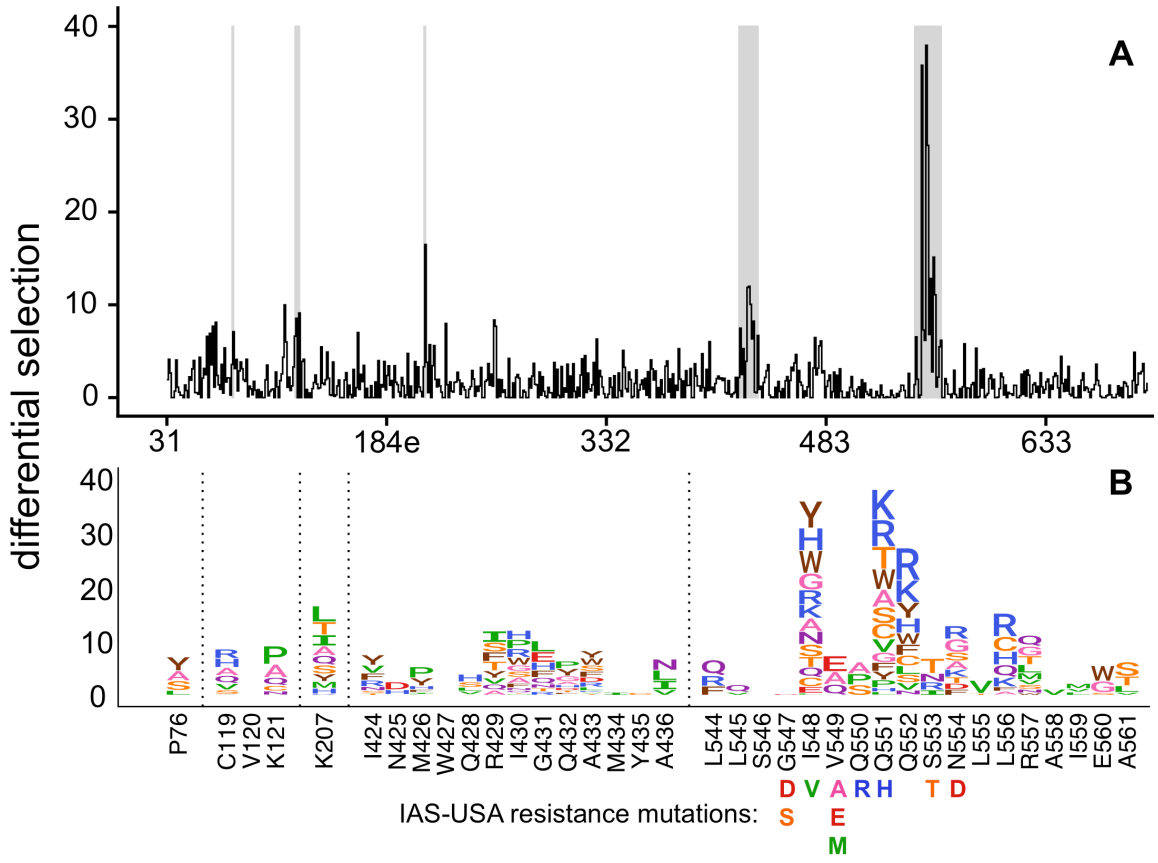


Figure 5.1. Complete map of enfuvirtide resistance.

A. The positive site differential selection is plotted across the mutagenized portion of Env. **B**. The mutation-level resistance profile for regions of interest (highlighted in grey in A). The height of each amino acid is proportional to its differential selection. All mutations from the 2017 IAS-USA enfuvirtide resistance mutations list [14] are labeled below their site.

We also identified resistance mutations in the NHR region that had not been previously characterized. For example, while Q551H is a known resistance mutation, our data revealed that many additional amino acids at this site also increased enfuvirtide resistance, with 12 other mutations at this site having a larger effect than Q551H (Figure 5.1). Further, mutations to

additional sites in NHR not included in the IAS–USA drug resistance mutations list, such as sites 552, 556, 557, and 560, also increased resistance to enfuvirtide (Figure 5.1).

Many of these resistance mutations cluster in the binding site of enfuvirtide in the NHR (Figure 5.2A). Site 551, where many different mutations confer resistance, directly interacts with enfuvirtide. Surprisingly, the side chains at the sites of the largest-effect mutations (548 and 552) face other NHR α-helices or the NHR trimer center (Figure 5.2A), and resistance at these sites occurred primarily due to mutations to positively charged or bulky amino acids (Figure 5.1B). The next two largest effect sites in this region, 554 and 556, appear to interact with enfuvirtide (Figure 5.2A).

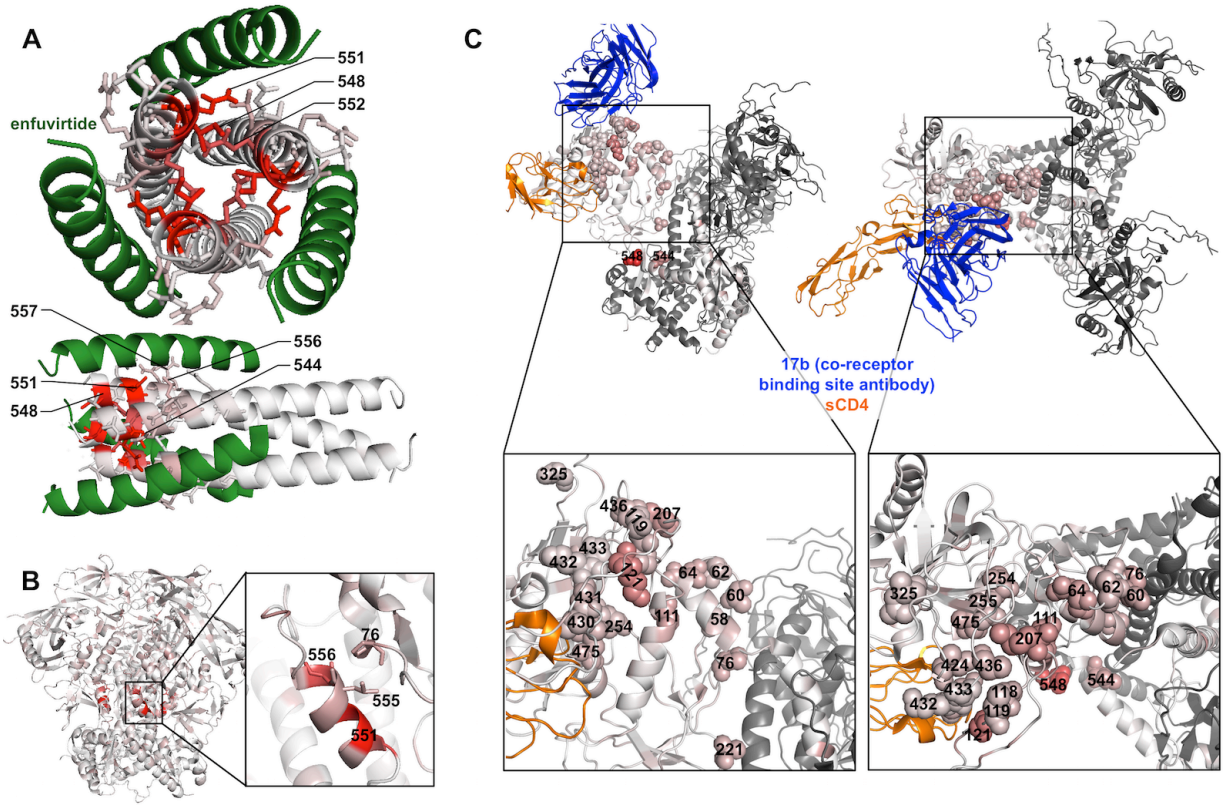


Figure 5.2. Locations of resistance mutations in different Env conformations.

A. A model of enfuvirtide (green) binding to the NHR coiled-coil trimer (colored according to the positive site differential selection at each site). The top 5% of all resistance sites are shown

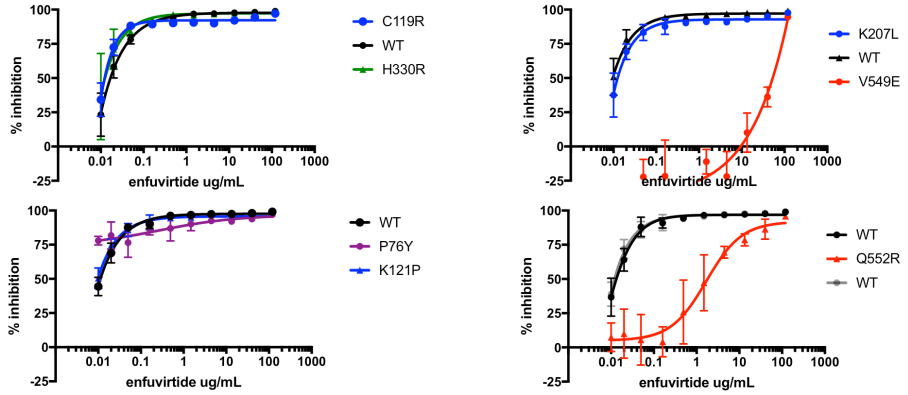
with sticks. While there is not a structure of enfuvirtide bound to Env, we used the post fusion, 6-helical bundle structure, composed of NHR and CHR peptides, to model enfuvirtide binding. This model assumes that enfuvirtide binding to NHR is similar to the CHR peptide binding to NHR in this structure. Portions of the CHR present in the structure that do not overlap with enfuvirtide are omitted, and the entirety of enfuvirtide is not present in the structure. PDB: 1AIK **B**. The closed, pre-fusion conformation of BG505 Env colored according to the maximum mutation differential at each site. PDB: 5YFK **C**. The side- and top-view of post-CD4-bound, open structure of BG505 Env, also bound by the co-receptor binding site antibody 17b, colored according to the maximum mutation differential at each site. The top 5% of resistance sites are shown with spheres. PDB: 5VN3.

There was also modest but reproducible enrichment of mutations at other Env sites outside of the NHR domain. One such mutation was P76Y, which interacts with NHR sites L555 and L556 near the enfuvirtide binding site in the prefusion conformation (Figure 2B). Other resistance mutations occurred at sites 424-436 in the β20/β21 strand of C4, as well as sites 119, 121, and 207 in the V1/V2 stem. While the V1/V2 stem is distant from β20/β21 in the prefusion Env conformation, it shifts upon CD4 binding to form the 4-stranded bridging sheet along with the β20/β21 strand to create the co-receptor binding site (Figure 2C). Similarly, additional sites of modest resistance mutations, including sites 60, 62, 64, 111, 254, 255, and 475 all cluster near the co-receptor binding site in the open, CD4-bound state (Figure 5.2C).

To validate our high-throughput mapping of enfuvirtide resistance mutations, we generated and tested individual BG505 Env pseudoviruses bearing single mutations for enfuvirtide sensitivity. In the NHR domain, both the previously characterized V549E mutation

and the Q552R mutation we identified increased resistance, shifting the IC_{50} by >150-fold (Figure 5.3). Other mutations that were modestly enriched (P76Y, C119R, K121P, and K207L) had little effect on IC_{50} but instead decreased the maximal inhibition plateau at high concentrations of drug, including the 8 ug/mL concentration used in our experiments (Supplementary Figure 3). Notably, both these validation experiments and the high-throughput mapping were performed with 100ug/mL DEAE-dextran. When the TZM-bl assays were repeated with just 10ug/mL DEAE-dextran, some of the resistance phenotypes were not as prominent (Figure 5.4).

A



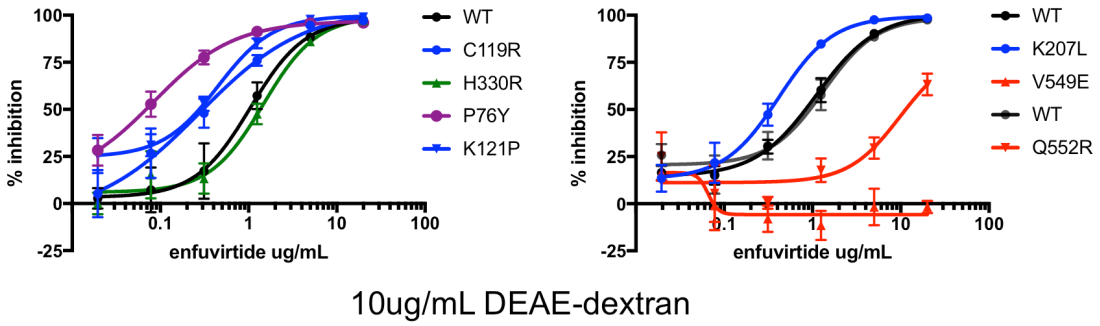
B

Mutational antigenic profiling — **TZM-bl inhibition assay (100ug/mL DEAE-dextran)**

	Mutant	mutdiffsel	IC50 (ug/mL)		Fold change in IC50		Maximum % inhibition	
			mean	standard error	mean	standard error	mean	standard error
plate 1	C119R	1.71	0.0124	0.0009	0.75	0.12	92.1	0.39
	WT	-	0.0168	0.0014	1.00	-	97.7	0.23
	H330R	-1.52	0.0120	0.0052	0.69	0.25	97.3	0.29
plate 2	WT	-	0.0113	0.0002	1.00	-	97.7	0.58
	P76Y	2.53	ambiguous	-	ambiguous	-	96.5	0.16
	K121P	3.37	0.0096	0.0007	0.86	0.08	95.8	0.57
plate 3	WT	-	0.0132	0.0015	1.00	-	97.0	0.09
	Q552R	5.93	1.8550	0.8930	134.20	52.27	89.8	2.74
	WT	-	0.0121	0.0011	0.92	0.02	96.9	0.13
plate 4	K207L	2.94	0.0120	0.0018	1.31	0.13	92.9	1.25
	WT	-	0.0094	0.0023	1.00	-	97.2	0.44
	V549E	2.89	51.3694	1.8818	5767.67	1217.06	ambiguous	-

Figure 5.3. Validation of enfuvirtide resistance mutants using a TZM-bl inhibition assay. TZM-bl inhibition assays were performed in the presence of 100ug/mL DEAE-dextran, similar to the resistance profiling. **A**. Inhibition curves are the average of two biological replicates, each performed in duplicate. **B**. The IC50, the fold change in IC50 relative to wildtype (WT), and the maximum percent inhibition for each mutant, determined from the fit four-parameter logistic curves. At least one WT virus was run on each plate and each mutant virus curve was compared to the plate internal WT control. Values were computed independently for each experimental replicate, then averaged across replicates. The standard error of the mean is also shown. H330R, which was not enriched in the resistance profiling, was included as a control. In A and B, mutant psuedoviruses are colored according to groups (black : WT; green : control mutant not expected to affect enfuvirtide sensitivity; blue : mutants in the V1/V2 Stem/co-receptor binding site; red : mutants in/near NHR binding site

A



10ug/mL DEAE-dextran

B

		Mutational antigenic profiling	TZM-bl inhibition assay (10ug/mL DEAE-dextran)					
			IC50 (ug/mL)		Fold change in IC50		Maximum % inhibition	
	Mutant	mutdiffsel	mean	standard error	mean	standard error	mean	standard error
plate 1	C119R	1.71	0.3035	0.07897	0.29	0.04	107.8	7.65
	H330R	-1.52	1.3316	0.13507	1.36	0.32	99.7	3.09
	P76Y	2.53	0.0697	0.00764	0.07	0.00	97.0	0.63
	K121P	3.37	0.2739	0.01859	0.27	0.02	100.0	0.34
	WT	-	1.0155	0.14005	1.00	-	98.7	0.04
plate 2	K207L	2.94	0.3377	0.02876	0.36	0.08	99.5	0.15
	V549E	2.89	>20	-	>25	-	24.5	6.86
	WT	-	0.9677	0.14323	1.00	-	98.2	0.88
	Q552R	5.93	10.8560	1.45797	11.24	0.16	76.4	9.22
	WT	-	0.8335	0.01385	0.88	0.12	100.7	1.18

Figure 5.4. TZM-bl inhibition assay performed with 10ug/mL DEAE-dextran instead of the 100ug/mL concentration used in Figure 5.3 and resistance profiling experiments. A. and **B**. As in Figure 5.3.

Discussion

We have quantified the effect of all single amino acid mutations to the extracellular and transmembrane ectodomain of BG505 Env on resistance to the fusion inhibitor enfuvirtide. This

map of resistance mutations includes both previously characterized and numerous novel resistance mutations. This dataset may be of use in the clinical monitoring of resistance during therapy and the genotypic prediction of enfuvirtide sensitivity prior to treatment.

The selected mutations also help elucidate Env's diverse mechanisms of enfuvirtide resistance, which likely include both disrupting enfuvirtide binding and altering fusion kinetics/conformations in multiple ways. While some mutations to NHR may directly disrupt interactions with enfuvirtide (e.g. site 551), other resistance mutations in this region appear to introduce positive charges or bulky amino acids at the center of the NHR coiled-coil. These mutations may slightly alter the coiled-coil structure to favor the intramolecular binding of the CHR domain over binding to enfuvirtide. It is possible that this dynamic is influenced by the subtype of the virus, as enfuvirtide is a mimetic of a subtype B CHR domain and differs from BG505 at 7 of 36 amino acids.

Other resistance mutations likely alter fusion kinetics or accessibility of the enfuvirtide binding site. P76Y may result in additional hydrophobic interactions between site 76 and L555 or L556 in the NHR in the prefusion conformation (Figure 5.2B). This interaction could directly limit access of enfuvirtide to the NHR binding site or delay the formation of the enfuvirtide-sensitive, CD4-induced conformational intermediate present post CD4-binding.

We also uncovered many small-effect size resistance mutations throughout the co-receptor binding site. It has been previously shown that co-receptor tropism, CCR5 affinity, and variability in V3 can affect enfuvirtide sensitivity [154,155], but we characterized resistance mutations at numerous additional sites in or near the co-receptor binding site, including sites in the V1/V2 loop stem and C4. Such mutations may increase CCR5 receptor affinity and/or fusion kinetics, limiting the time in which a CD4-induced, enfuvirtide sensitive conformational

intermediate is exposed. However, there is also evidence of weak interactions between enfuvirtide and the co-receptor binding site on gp120 [157]; resistance mutations in this region may disrupt this potential secondary enfuvirtide-Env interaction.

The complex mechanisms of enfuvirtide resistance were highlighted further when validating these resistance mutations with traditional inhibition curves. While mutations in the NHR domain resulted in a shift in the inhibition curve, the remaining mutants we tested resulted in a decrease in the maximum percent inhibition rather than a shift in the inhibition curve. It remains to be determined if such phenotypes manifest *in vivo*. However, if a small sub-population of viruses were resistant to enfuvirtide at *in vivo* drug concentrations, this could present an opportunity for additional resistance to enfuvirtide or other drugs in combination therapies to arise.

While the resistance mutations uncovered here should be considered when evaluating clinical resistance, they also come with some caveats. We examined the effect of single amino acids in a single subtype A Env. Many patients receiving enfuvirtide have subtype B infections, and there could be strain-specific differences in the effects of mutations on enfuvirtide resistance. For example, we do not identify the G547D and G547S resistance mutations, which were identified by enfuvirtide selection of lab-adapted, subtype B viruses in cell culture [151]. Our previous studies show that these mutations are tolerated in a lab-adapted subtype B virus, but they are not well-tolerated by the transmitted subtype A variant BG505 during viral replication in cell culture [51,52]. These differences highlight how the balance between enfuvirtide resistance and replicative fitness may depend on the virus strain.

In summary, this map of enfuvirtide resistance identifies numerous previously uncharacterized and potentially clinically relevant resistance mutations, highlighting the utility of

using viral deep mutational scanning approaches to evaluate drug resistance. This data also sheds light on the mechanisms of enfuvirtide resistance and Env's fusion process.

Chapter VI

Conclusions and future directions

The ability to comprehensively assay the phenotypic effects of all mutations to HIV-1 Env is just starting to realize it's potential. We have leveraged mutational antigenic profiling to study topics related to antibody escape, and have started to refine vaccine design approaches, evaluate antibody immunotherapies, and profile drug resistance. Future projects could build upon these studies and expand into additional questions. I speculate on some of the promising future directions in the following subsections.

Mapping the development and determinants of antibody breadth

Defining the determinants of antibody breadth, and the virological and immunological processes that trigger the development of antibody breadth, is crucial to rationally designing vaccines in hope of eliciting similarly broad responses.

It is well known that bnAbs target functionally conserved regions, and there is ample intuitive and experimental evidence that targeting more conserved regions results in greater breadth. For example, the CH235 bnAb lineage increased its precision in binding the CD4bs as the lineage developed breadth, and among all CD4bs bnAbs, this precision of binding inside of vs. outside of the CD4bs correlates with breadth [158]. However, there is still considerable diversity with structurally-defined bnAb epitopes (reviewed in [159]). Thus, it is easy to make the argument that breadth could also driven by the ability to tolerate diversity within the structural epitope.

There is also considerable evidence on how viral diversity drives antibody affinity maturation and greater antibody breadth. *In vivo* viral genetic diversity within a bnAb-targeted epitope has been shown to precede or coincide the development of breadth [160,161], suggesting this viral diversity could trigger affinity maturation that allows an antibody to tolerate diversity within the epitope. This concept was furthered by Bhiman and colleagues [162] who showed, among 33 clonal variants of the CAP256-VRC26 antibody lineage, the ability to neutralize autologous pseudoviruses bearing the minority epitope variants that arose *in vivo*, termed "immunotypes", correlated strongly with the breadth of the antibody, and this phenotype clustered on the antibody clade that went on to develop the greatest breadth. Strikingly, this same pattern of epitope diversification also occurred in another patient who developed a similar V2/apex bnAb response (PCT64) [163]. Together, these studies suggest there are reproducible patterns of viral evolution that drive and/or are driven by the development of bnAb breadth, which could possibly be mimicked with vaccination strategies.

We can leverage mutational antigenic profiling to quantify the effect of all single amino acid variants (i.e., many more immunotypes). When applied to reconstructed or longitudinally sampled members of an antibody lineage, we can examine how the antigenic effects of all mutations change as an antibody develops breadth. As an antibody develops breadth, how does the functional epitope change? For example, are there fewer amino-acid mutations at key epitope sites that mediate escape? Or are there fewer distinct regions of escape? These questions can be addressed by using mutational antigenic profiling to map escape from multiple monoclonal antibodies from a single developmental lineage.

Further, sequencing the viral population from patient samples can allude to how viral diversity drove these changes and vice versa. Does an antibody "learn" to neutralize particular

epitope variants only after they have arisen *in vivo* and been presented to the evolving B-cells, as suggested by Bhiman *et al.* [164]? How many "immunotypes" play a role in this process? How does this subsequent response drive further viral evolution?

We have *started* to address portions of these questions by profiling escape from two clonal variants, the anti-N332 glycan supersite bnAbs PGT121 and 10-1074 [54]. We found interesting differences between the two antibodies in the specific mutations that mediated escape, as well as the overall effect sizes of escape mutations at key sites (Figure 3.6A,B). For instance, mutations to site 325 had a larger effect on 10-1074 than PGT121, as did mutations that disrupted the N332 PNG. The basis for differences in escape between PGT121 and 10-1074 may be explained by their differential somatic hypermutation. There are numerous differences in the portions of these two antibodies that contact the N332 glycan (Figure 6.1), and these differences have previously been shown to affect glycan recognition [101]. In contrast, antibody residues that directly contact site 325 are conserved between PGT121 and 10-1074. However, there are numerous differences in the light chain variable loops near these contact sites that may be responsible for the differential effects of mutations at site 325 on the two antibodies (Figure 6.1).

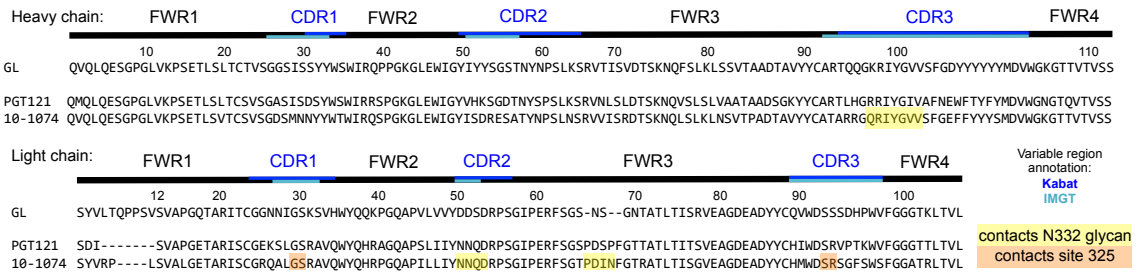


Figure 6.1 Differences in SHM of PGT121 and 10-1074 may explain differences in viral escape. The heavy (top) and light chain (bottom) genes of the clonal variants PGT121 and 10-1074 are aligned, along with their inferred germline (GL). Kabat and IMGT variable regions

annotations are shown. There are several sites in Env where the effects of mutations differ markedly between PGT121 and 10-1074, most prominently sites 325 and the N332 PNG (Figure 3.6A,B). We used the 10-1074 / Env co-crystal structure 5TZ3 to identify sites in 10-1074 that contact these Env sites, and have highlighted them in the alignment above PGT121 contact sites are not labeled, as there is not a high-resolution structure of available. The germline inference, alignment, and variable loop annotation are from Mouquet *et al.* 2012 [101], which also discusses additional structural differences between these antibodies (such as differences in the light chain loops of the unliganded antibody structures).

However, PGT121 and 10-1074 are only two distant members of an antibody lineage, both of which developed considerable breadth (69% and 67% breadth). In contrast, mapping differences in escape of clonal variants as breadth develops could elucidate the basis for the development of breadth. Fortunately, colleagues in the Overbaugh and Matsen groups have reconstructed the developmental lineage of and infant N332 glycan targeting bnAb BF520.1 and characterized the breadth of intermediate antibodies [165]. Next, we will profile how viral escape changes as this antibody lineage develops breadth. Further, we will deep sequence the viral population from longitudinal timepoints to examine how well our escape profiles predict viral evolution *in vivo.* This will also help to disentangle the interplay between viral evolution and the development of antibody breadth.

Excitingly, the Overbaugh lab has also isolated the QA013.2 bnAb from superinfected donor that targets the same N332 glycan epitope as BF520.1 [166]. It is of interest to examine the development of antibody breadth for a similar bnAb in another patient. Further, do similar patterns of N332 glycan supersite epitope evolution occur in QA013 and BF520, similar to the

parallel patterns of V2/apex epitope evolution that drove the CAP256-VRC26 (which also occurred after superinfection) and PCT64 antibody lineages?

Modeling viral escape from antibodies and its biochemical basis

Observational studies of how well mutational antigenic profiling predicts viral evolution *in vivo* leads to bolder goals: predicting viral escape from a given antibody of interest. The high-dimensional functional and antigenic datasets we have generated could have utility in computationally modeling viral escape, the interplay between immune escape and viral fitness, and even interrogating the biochemical basis of evading antibody responses.

We have previously made high-throughput measurements of how well each amino-acid mutation to Env is tolerated for viral replication in cell culture [51,52]. These estimates quantify the each mutational effect of each mutation, a proxy for fitness. In the simplest form, we can we can examine how the mutational tolerance influences the ability of mutations to escape each antibody. Similarly, we can examine the epitope sequence variation in nature in parallel. Example data illustrating this simple visual analysis for 10E8 is shown if Figure 6.2.

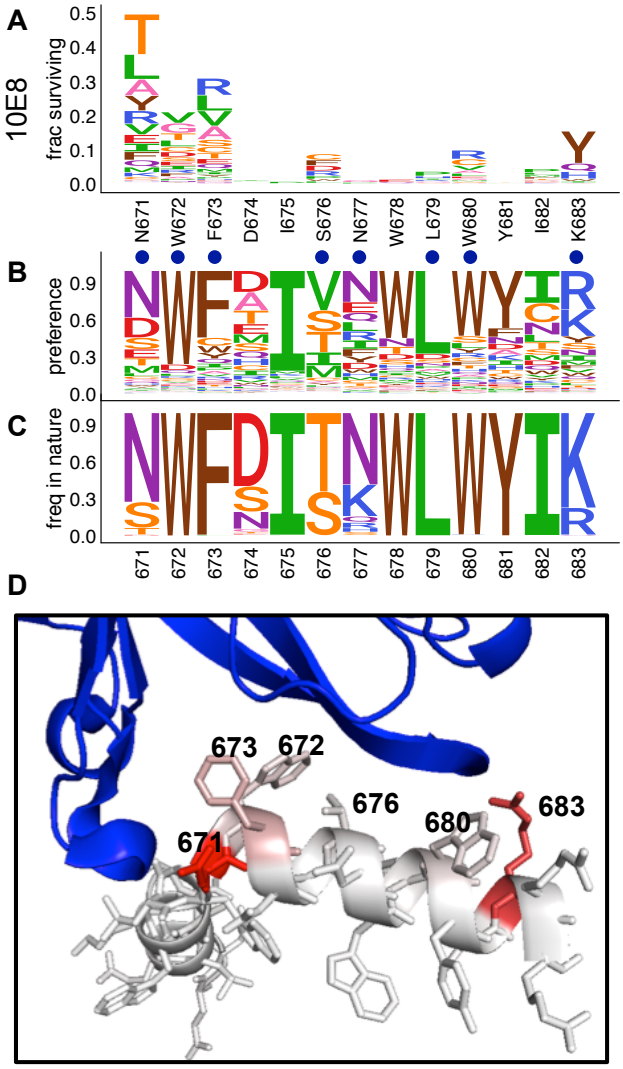


Figure 6.2 The escape profile, mutational tolerance, and natural sequence variation of the 10E8 epitope. A. The excess fraction surviving 10E8. Blue spheres indicate structurally defined contacts. **B**. The BG505 mutational tolerance of the 10E8 epitope. The height of each amino acid corresponds to how well tolerated it is for viral replication in cell culture. Data from [52]. **C**. The natural sequence variation within the epitope, from LANL's Group M filtered web alignment [123]. **D**. 10E8 is shown in blue, and the MPER peptide is colored according to the maximum excess fraction surviving at each site (PDB 4G6F). As in Figure 3.8E.

Including these measures of antigenicity, mutational tolerance, and natural sequence variation in phylogenetic and/or biophysical models of viral evolution may enable more precise modeling of viral escape from antibodies. Such models could be used to predict viral escape, determine the hardest antibody to escape from, as well as to evaluate viral evolution in clinical trials of antibody immunotherapies.

Further, combining such models or these datasets with structural bioinformatics could interrogate the biochemical basis of viral escape. This could include quantifying the contribution of each residue/interaction to the antibody's binding energetics or structurally modeling the consequence of each mutation. Together, these approaches could reveal patterns in the physiochemical properties within and near to the structural epitope that allow Env to evade bnAbs.

Genotypic prediction of phenotypes

Predicting antibody (or drug) sensitivity based on genotype is important for sieve analyses of vaccines and therapies, personalized use of bnAb immunotherapies, and the targeted use of specific antiretroviral drugs in different populations. Our high-throughput measurements of the phenotypic effect of all possible mutations are exemplary datasets to use in predicting or scoring genotypes for the sensitivity to an antibody or drug, as has been preciously done for drug sensitivity [84]. Assaying phenotypes across multiple strains will be imperative to refining any broadly applicable genotype-based phenotype prediction algorithm.

Scaling mutational antigenic profiling to broadly impact vaccines and therapies

While mutational antigenic profiling has proven powerful, it is still limited by throughput, cost, and experimental constraints that only allow us to measure single amino-acid mutations to a single viral strain. For mutational antigenic profiling to broadly impact HIV vaccines and therapies, its throughput must be scaled. Further, these experimental approaches, and the high-dimensional data they generate, must be made easily accessible to labs across the field.

If throughput can be increased substantially, the number of questions possible to address would scale accordingly. Briefly, this could include mapping escape from different pools of bnAbs in order to experimentally select the bnAb combination that is hardest to escape from. Further, it could be leveraged to map the multiple components of polyclonal antibody responses, both in natural infection and in evaluating vaccine responses. Mapping escape from the numerous (>30) members of well-characterized antibody lineages, such as CH235 and CAP256-VRC26, could lend further insights into the determinants of antibody breadth. Further, all of these analyses would be greatly strengthened by examining the antigenic effects of mutations across multiple viral strains, as well as testing the effect of multiple amino-acid mutations.

Conclusion

In conclusion, HIV Env mutational antigenic profile has yielded unprecedented views of the functional interfaces between antibodies and Env. This has allowed for the unbiased definition the functional epitope of bnAbs for the first time, which are distinct from their

structurally defined epitopes. While we have leveraged this approach to refine and evaluate vaccine design, antibody immunotherapies, and drug resistance, it's full potential will not be realized until it is widely applied to throughout the field.

BIBLIOGRAPHY

1 UNAIDS. UNAIDS 2018 Global HIV Statistics(2017). 2018; :1–5.

2 Muessig KE, Cohen MS. Advances in HIV Prevention for Serodiscordant Couples. *Curr HIV/AIDS Rep* 2014; **11**:434–446.

3 Korber B, Gaschen B, Yusim K, Thakallapally R, Kesmir C, Detours V. Evolutionary and immunological implications of contemporary HIV-1 variation. *Br Med Bull* 2001; **58**:19–42.

4 Ozorowski G, Pallesen J, Val N De, Lyumkis D, Cottrell CA, Torres JL, *et al.* Open and closed structures reveal allostery and pliability in the HIV-1 envelope spike. *Nat Publ Gr* Published Online First: 2017. doi:10.1038/nature23010

5 Wilen CB, Tilton JC, Doms RW, Walker B, Mcmichael A, Wilen CB, *et al.* HIV : Cell Binding and Entry. 2012; :1–14.

6 Chan DC, Kim PS. HIV entry and its inhibition. *Cell* 1998; **93**:681–684.

7 Stewart-Jones GBE, Soto C, Lemmin T, Chuang GY, Druz A, Kong R, *et al.* Trimeric HIV-1-Env Structures Define Glycan Shields from Clades A, B, and G. *Cell* 2016; **165**:813–826.

8 Klein JS, Bjorkman PJ. Few and far between: How HIV may be evading antibody avidity. *PLoS Pathog* 2010; **6**:1–6.

9 Munro JB, Gorman J, Ma X, Zhou Z, Arthos J, Burton DR, *et al.* Conformational dynamics of single HIV-1 envelope trimers on the surface of native virions. *Science (80-)* 2014; **346**:759–763.

10 Pancera M, Zhou T, Druz A, Georgiev IS, Soto C, Gorman J, *et al.* Structure and immune recognition of trimeric pre-fusion HIV-1 Env. *Nature* 2014; **514**:455–461.

11 Burton DR, Mascola JR. Antibody responses to envelope glycoproteins in HIV-1 infection. *Nat Immunol* 2015; **16**:571–576.

12 Overbaugh J, Morris L. The antibody response against HIV-1. *Cold Spring Harb Perspect Med* 2012; **2**:1–18.

13 Keele BF, Giorgi EE, Salazar-Gonzalez JF, Decker JM, Pham KT, Salazar MG, *et al.* Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proc Natl Acad Sci* Published Online First: 2008. doi:10.1073/pnas.0802203105

14 Wei X, Decker JM, Wang S, Hui H, Kappes JC, Wu X, *et al.* Antibody neutralization and escape by HIV-1. *Nature* 2003; **422**:307–312.

15 Richman DD, Wrin T, Little SJ, Petropoulos CJ. Rapid evolution of the neutralizing antibody response to HIV type 1 infection. *Proc Natl Acad Sci U S A* 2003; **100**:4144–9.

16 Sagar M, Wu X, Lee S, Overbaugh J. Human immunodeficiency virus type 1 V1-V2 envelope loop sequences expand and add glycosylation sites over the course of infection, and these modifications affect antibody neutralization sensitivity. *J Virol* 2006; **80**:9586–98.

17 Ronen K, Dingens AS, Graham SM, Jaoko W, Mandaliya K, McClelland RS, *et al.* Comprehensive Characterization of Humoral Correlates of Human Immunodeficiency Virus 1 Superinfection Acquisition in High-risk Kenyan Women. *EBioMedicine* 2017; **18**:216–224.

18 Wu X, Kong X-P. Antigenic landscape of the HIV-1 envelope and new immunological concepts defined by HIV-1 broadly neutralizing antibodies. *Curr Opin Immunol* 2016; **42**:56–64.

19 Bonsignori M, Liao H-X, Gao F, Williams WB, Alam SM, Montefiori DC, *et al.* Antibody-virus co-evolution in HIV infection: paths for HIV vaccine development. *Immunol Rev* 2017; **275**:145–160.

20 Lee JH, Andrabi R, Su CY, Yasmeen A, Julien JP, Kong L, *et al.* A Broadly Neutralizing Antibody Targets the Dynamic HIV Envelope Trimer Apex via a Long, Rigidified, and Anionic β-Hairpin Structure. *Immunity* 2017; **46**:690–702.

21 Zhou T, Lynch RM, Chen L, Acharya P, Wu X, Doria-Rose N a., *et al.* Structural repertoire of HIV-1-neutralizing antibodies targeting the CD4 supersite in 14 donors. *Cell* 2015; **161**:1280–1292.

22 de Taeye SW, Moore JP, Sanders RW. HIV-1 Envelope Trimer Design and Immunization Strategies To Induce Broadly Neutralizing Antibodies. *Trends Immunol* 2016; **37**:221–232.

23 Steichen JM, Kulp DW, Tokatlian T, Escolano A, Dosenovic P, Stanfield RL, *et al.* HIV Vaccine Design to Target Germline Precursors of Glycan-Dependent Broadly Neutralizing Antibodies. *Immunity* 2016; **45**:483–496.

24 Escolano A, Steichen JM, Dosenovic P, Kulp DW, Golijanin J, Sok D, *et al.* Sequential Immunization Elicits Broadly Neutralizing Anti-HIV-1 Antibodies in Ig Knockin Mice. *Cell* 2016; **166**:1445–1458.e12.

25 Kong R, Xu K, Zhou T, Acharya P, Lemmin T, Liu K, *et al.* Fusion peptide of HIV-1 as a site of vulnerability to neutralizing antibody. *Science* 2016; **352**:828–33.

26 Xu K, Acharya P, Kong R, Cheng C, Chuang G-Y, Liu K, *et al.* Epitope-based vaccine design yields fusion peptide-directed antibodies that neutralize diverse strains of HIV-1. *Nat Med* 2018; **24**:857–867.

27 Sok D, Briney B, Jardine JG, Kulp DW, Menis S, Wood A, *et al.* Priming HIV-1 broadly neutralizing antibody precursors in human Ig loci transgenic mice. *Science (80-)* 2016; **3945**:1–10.

28 Briney B, Sok D, Jardine JG, Kulp DW, Skog P, Menis S, *et al.* Tailored Immunogens Direct Affinity Maturation toward HIV Neutralizing Antibodies. *Cell* 2016; **166**:1459–1470.e11.

29 Hessell AJ, Poignard P, Hunter M, Hangartner L, Tehrani DM, Bleeker WK, *et al.* Effective, low-titer antibody protection against low-dose repeated mucosal SHIV challenge in macaques. *Nat Med* 2009; **15**:951–4.

30 Mascola JR, Lewis MG, Stiegler G, Harris D, VanCott TC, Hayes D, *et al.* Protection of Macaques against pathogenic simian/human immunodeficiency virus 89.6PD by passive transfer of neutralizing antibodies. *J Virol* 1999; **73**:4009–18.

31 Mascola JR, Stiegler G, VanCott TC, Katinger H, Carpenter CB, Hanson CE, *et al.* Protection of macaques against vaginal transmission of a pathogenic HIV-1/SIV chimeric virus by passive infusion of neutralizing antibodies. *Nat Med* 2000; **6**:207–210.

32 Moldt B, Rakasz EG, Schultz N, Chan-Hui P-Y, Swiderek K, Weisgrau KL, *et al.* Highly potent HIV-specific antibody neutralization in vitro translates into effective protection against mucosal SHIV challenge in vivo. *Proc Natl Acad Sci U S A* 2012; **109**:18921–5.

33 Parren PWHI, Marx PA, Hessell AJ, Luckay A, Harouse J, Cheng-Mayer C, *et al.* Antibody Protects Macaques against Vaginal Challenge with a Pathogenic R5 Simian/Human Immunodeficiency Virus at Serum Levels Giving Complete Neutralization In Vitro. *J Virol* 2001; **75**:8340–8347.

34 Pegu A, Yang Z -y., Boyington JC, Wu L, Ko S-Y, Schmidt SD, *et al.* Neutralizing antibodies to HIV-1 envelope protect more effectively in vivo than those to the CD4 receptor. *Sci Transl Med* 2014; **6**:243ra88-243ra88.

35 Xu L, Xu L, Pegu A, Rao E, Doria-rose N, Beninga J, *et al.* Trispecific broadly neutralizing HIV antibodies mediate potent SHIV protection in macaques. 2017; **8630**:1–13.

36 Barouch DH, Whitney JB, Moldt B, Klein F, Oliveira TY, Liu J, *et al.* Therapeutic efficacy of potent neutralizing HIV-1-specific monoclonal antibodies in SHIV-infected rhesus monkeys. *Nature* 2013; **503**:224–8.

37 Horwitz J a., Halper-Stromberg a., Mouquet H, Gitlin a. D, Tretiakova a., Eisenreich TR, *et al.* HIV-1 suppression and durable control by combining single broadly neutralizing antibodies and antiretroviral drugs in humanized mice. *Proc Natl Acad Sci* 2013; **110**:16538–16543.

38 Klein F, Halper-Stromberg A, Horwitz J a., Gruell H, Scheid JF, Bournazos S, *et al.* HIV therapy by a combination of broadly neutralizing antibodies in humanized mice. *Nature*

2012; **492**:118–122.

39 Shingai M, Nishimura Y, Klein F, Mouquet H, Donau OK, Plishka R, *et al.* Antibody-mediated immunotherapy of macaques chronically infected with SHIV suppresses viraemia. *Nature* 2013; **503**:277–80.

40 Margolis DM, Koup RA, Ferrari G. HIV antibodies for treatment of HIV infection. *Immunol Rev* 2017; **275**:313–323.

41 Hessell AJ, Jaworski JP, Epson E, Matsuda K, Pandey S, Kahl C, *et al.* Early short-term treatment with neutralizing human monoclonal antibodies halts SHIV infection in infant macaques. *Nat Med* 2016; **22**:362–8.

42 Borducchi EN, Liu J, Nkolola JP, Cadena AM, Yu W-H, Fischinger S, *et al.* Antibody and TLR7 agonist delay viral rebound in SHIV-infected monkeys. *Nature* 2018; **563**:360–364.

43 Caskey M, Klein F, Lorenzi JCC, Seaman MS, West AP, Buckley N, *et al.* Viraemia suppressed in HIV-1-infected humans by broadly neutralizing antibody 3BNC117. *Nature* 2015; **522**:487–491.

44 Bar KJ, Sneller MC, Harrison LJ, Justement JS, Overton ET, Petrone ME, *et al.* Effect of HIV Antibody VRC01 on Viral Rebound after Treatment Interruption. *N Engl J Med* 2016; :NEJMoa1608243.

45 Lynch RM, Boritz E, Coates EE, DeZure A, Madden P, Costner P, *et al.* Virologic effects of broadly neutralizing antibody VRC01 administration during chronic HIV-1 infection. *Sci Transl Med* 2015; **7**:319ra206-319ra206.

46 Caskey M, Schoofs T, Gruell H, Settler A, Karagounis T, Kreider EF, *et al.* Antibody 10-1074 suppresses viremia in HIV-1-infected individuals. *Nat Med* 2017; **23**:185–191.

47 Scheid JF, Horwitz JA, Bar-On Y, Kreider EF, Lu C-L, Lorenzi JCC, *et al.* HIV-1 antibody 3BNC117 suppresses viral rebound in humans during treatment interruption. *Nature* 2016; **535**:556–560.

48 Gruell H, Klein F. Antibody-mediated prevention and treatment of HIV-1 infection. *Retrovirology* 2018; **15**:73.

49 Poignard P, Sabbe R, Picchio GR, Wang M, Gulizia RJ, Katinger H, *et al.* Neutralizing antibodies have limited effects on the control of established HIV-1 infection in vivo. *Immunity* 1999; **10**:431–438.

50 Trkola A, Kuster H, Rusert P, Joos B, Fischer M, Leemann C, *et al.* Delay of HIV-1 rebound after cessation of antiretroviral therapy through passive transfer of human neutralizing antibodies. *Nat Med* 2005; **11**:615–622.

51 Haddox HK, Dingens AS, Bloom JD. Experimental Estimation of the Effects of All Amino-Acid Mutations to HIV's Envelope Protein on Viral Replication in Cell Culture.

PLOS Pathog 2016; **12**:e1006114.

52 Haddox HK, Dingens AS, Hilton SK, Overbaugh J, Bloom JD. Mapping mutational effects along the evolutionary landscape of HIV envelope. *Elife* 2018; **7**:e34420.

53 Dingens AS, Haddox HK, Overbaugh J, Bloom JD. Comprehensive Mapping of HIV-1 Escape from a Broadly Neutralizing Antibody. *Cell Host Microbe* 2017; **21**:777–787.e4.

54 Dingens AS, Arenz D, Weight H, Overbaugh J, Bloom JD. An antigenic atlas of HIV-1 escape from broadly neutralizing antibodies. *bioRxiv* Published Online First: 2018. doi:doi.org/10.1101/406355

55 Dingens AS, Acharya P, Haddox HK, Rawi R, Xu K, Chuang G-Y, *et al.* Complete functional mapping of infection- and vaccine-elicited antibodies against the fusion peptide of HIV. *PLOS Pathog* 2018; **14**:e1007159.

56 Dingens AS, Arenz D, Overbaugh J, Bloom JD. Massively parallel profiling of HIV-1 resistance to the fusion inhibitor enfuvirtide. *bioRxiv* Published Online First: 2018. doi:10.1101/472746

57 Cunningham BC, Wells J a. Comparison of a structural and a functional epitope. J Mol Biol. 1993; **234**:554–563.

58 Clackson T, Wells J a. A hot spot of binding energy in a hormone-receptor interface. *Science (80-)* 1995; **267**:383–386.

59 Li Y, O'Dell S, Walker LM, Wu X, Guenaga J, Feng Y, *et al.* Mechanism of Neutralization by the Broadly Neutralizing HIV-1 Monoclonal Antibody VRC01. *J Virol* 2011; **85**:8954–8967.

60 Falkowska E, Ramos a., Feng Y, Zhou T, Moquin S, Walker LM, *et al.* PGV04, an HIV-1 gp120 CD4 Binding Site Antibody, Is Broad and Potent in Neutralization but Does Not Induce Conformational Changes Characteristic of CD4. *J Virol* 2012; **86**:4394–4403.

61 Doud MB, Hensley SE, Bloom JD. Complete mapping of viral escape from neutralizing antibodies. *PLOS Pathog* 2017; **13**:e1006271.

62 Jardine JG, Kulp DW, Havenar-Daughton C, Sarkar A, Briney B, Sok D, *et al.* HIV-1 broadly neutralizing antibody precursor B cells revealed by germline-targeting immunogen. *Science (80-)* 2016; **351**:1458–1463.

63 Wu X, Parast AB, Richardson BA, Nduati R, John-Stewart G, Mbori-Ngacha D, *et al.* Neutralization Escape Variants of Human Immunodeficiency Virus Type 1 Are Transmitted from Mother to Infant. 2006; **80**:835–844.

64 Goo L, Milligan C, Simonich C a., Nduati R, Overbaugh J. Neutralizing Antibody Escape during HIV-1 Mother-to-Child Transmission Involves Conformational Masking of Distal Epitopes in Envelope. *J Virol* 2012; **86**:9566–9582.

65 Simonich CA, Williams KL, Verkerke HP, Williams JA, Nduati R, Lee KK, *et al.* HIV-1 Neutralizing Antibodies with Limited Hypermutation from an Infant. *Cell* 2016; **166**:77–87.

66 Chakrabarti L, Emerman M, Tiollais P, Sonigo P. The cytoplasmic domain of simian immunodeficiency virus transmembrane protein modulates infectivity. *J Virol* 1989; **63**:4395–4403.

67 Yuste E, Reeves JD, Doms RW, Desrosiers RC. Modulation of Env Content in Virions of Simian Immunodeficiency Virus : Correlation with Cell Surface Expression and Virion Infectivity Modulation of Env Content in Virions of Simian Immunodeficiency Virus : Correlation with Cell Surface Expression and Vi. 2004; **78**:6775–6785.

68 Li Y, Luo L, Thomas DY, Kang OY. Control of Expression, Glycosylation, and Secretion of HIV-1 gp120 by Homologous and Heterologous Signal Sequences. *Virology* 1994; **204**:266–278.

69 Doud M, Bloom J. Accurate Measurement of the Effects of All Amino-Acid Mutations on Influenza Hemagglutinin. *Viruses* 2016; **8**:155.

70 Wu NC, Du Y, Le S, Young AP, Zhang T-H, Wang Y, *et al.* Coupling high-throughput genetics with phylogenetic information reveals an epistatic interaction on the influenza A virus M segment. *BMC Genomics* 2016; **17**:46.

71 Jabara CB, Jones CD, Roach J, Anderson J a., Swanstrom R. Accurate sampling and deep sequencing of the HIV-1 protease gene using a Primer ID. *Proc Natl Acad Sci* 2011; **108**:20166–20171.

72 Blattner C, Lee JH, Sliepen K, Derking R, Falkowska E, de la Peña AT, *et al.* Structural Delineation of a Quaternary, Cleavage-Dependent Epitope at the gp41-gp120 Interface on Intact HIV-1 Env Trimers. *Immunity* 2014; **40**:669–680.

73 Lee JH, Ozorowski G, Ward AB. Cryo-EM structure of a native, fully glycosylated, cleaved HIV-1 envelope trimer. *Science (80-)* 2016; **351**:1043–1048.

74 Falkowska E, Le KM, Ramos A, Doores KJ, Lee JH, Blattner C, *et al.* Broadly Neutralizing HIV Antibodies Define a Glycan-Dependent Epitope on the Prefusion Conformation of gp41 on Cleaved Envelope Trimers. *Immunity* 2014; **40**:657–668.

75 Van Gils MJ, Van Den Kerkhof TLGM, Ozorowski G, Cottrell C a, Sok D, Pauthner M, *et al.* An HIV-1 antibody from an elite neutralizer implicates the fusion peptide as a site of vulnerability. *Nat Microbiol* 2016; :1–10.

76 Wibmer CK, Gorman J, Ozorowski G, Bhiman JN, Sheward DJ, Elliott DH, *et al.* Structure and Recognition of a Novel HIV-1 gp120-gp41 Interface Antibody that Caused MPER Exposure through Viral Escape. *PLoS Pathog* 2017; **13**:e1006074.

77 McCoy LE, Falkowska E, Doores KJ, Le K, Sok D, van Gils MJ, *et al.* Incomplete

Neutralization and Deviation from Sigmoidal Neutralization Curves for HIV Broadly Neutralizing Monoclonal Antibodies. *PLoS Pathog* 2015; **11**:1–19.

78 Todd C a., Greene KM, Yu X, Ozaki D a., Gao H, Huang Y, *et al.* Development and implementation of an international proficiency testing program for a neutralizing antibody assay for HIV-1 in TZM-bl cells. *J Immunol Methods* 2012; **375**:57–67.

79 Diskin R, Klein F, Horwitz J a., Halper-Stromberg a., Sather DN, Marcovecchio PM, *et al.* Restricting HIV-1 pathways for escape using rationally designed anti-HIV-1 antibodies. *J Exp Med* 2013; **210**:1235–1249.

80 Lynch RM, Wong P, Tran L, O'Dell S, Nason MC, Li Y, *et al.* HIV-1 Fitness Cost Associated with Escape from the VRC01 Class of CD4 Binding Site Neutralizing Antibodies. *J Virol* 2015; **89**:4201–13.

81 Diskin R, Scheid JF, Marcovecchio PM, West AP, Klein F, Gao H, *et al.* Increasing the Potency and Breadth of an HIV Antibody by Using Structure-Based Rational Design. *Science (80-)* 2011; **334**:1289–1293.

82 Tural C, Ruiz L, Holtzer C, Schapiro J, Viciana P, González J, *et al.* Clinical utility of HIV-1 genotyping and expert advice: the Havana trial. *AIDS* 2002; **16**:209–218.

83 Lehman DA, Baeten JM, McCoy CO, Weis JF, Peterson D, Mbara G, *et al.* Risk of drug resistance among persons acquiring HIV within a randomized clinical trial of single-or dual-agent preexposure prophylaxis. *J Infect Dis* 2015; **211**:1211–1218.

84 Vercauteren J, Vandamme AM. Algorithms for the interpretation of HIV-1 genotypic drug resistance information. *Antiviral Res* 2006; **71**:335–342.

85 Poss M, Overbaugh J. Variants from the diverse virus population identified at seroconversion of a clade A human immunodeficiency virus type 1-infected woman have distinct biological properties. *J Virol* 1999; **73**:5255–5264.

86 Provine NM, Puryear WB, Wu X, Overbaugh J, Haigwood NL. The infectious molecular clone and pseudotyped virus models of human immunodeficiency virus type 1 exhibit significant differences in virion composition with only moderate differences in infectivity and inhibition sensitivity. *J Virol* 2009; **83**:9002–7.

87 Bloom JD. An experimentally determined evolutionary model dramatically improves phylogenetic fit. *Mol Biol Evol* 2014; **31**:1956–1978.

88 Korber BT, Foley BT, Kuiken CL, Pillai SK, Sodroski JG. Numbering positions in HIV relative to HXB2CG. *Hum retroviruses AIDS* 1998; **3**:102–111.

89 Benki S, McClelland RS, Emery S, Baeten JM, Richardson BA, Lavreys L, *et al.* Quantification of genital human immunodeficiency virus type 1 (HIV-1) DNA in specimens from women with low plasma HIV-1 RNA levels typical of HIV-1 nontransmitters. *J Clin Microbiol* 2006; **44**:4357–4362.

90 Strain MC, Lada SM, Luong T, Rought SE, Gianella S, Terry VH, *et al.* Highly Precise Measurement of HIV DNA by Droplet Digital PCR. *PLoS One* 2013; **8**:2–9.

91 Cortez V, Wang B, Dingens A, Chen MM, Ronen K, Georgiev IS, *et al.* The Broad Neutralizing Antibody Responses after HIV-1 Superinfection Are Not Dominated by Antibodies Directed to Epitopes Common in Single Infection. *PLOS Pathog* 2015; **11**:e1004973.

92 Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci U S A* 2001; **98**:10037–41.

93 Crooks GE, Hon G, Chandonia J-M, Brenner SE. WebLogo: a sequence logo generator. *Genome Res* 2004; **14**:1188–90.

94 Bloom JD. Software for the analysis and visualization of deep mutational scanning data. *BMC Bioinformatics* 2015; **16**:1–13.

95 Kwong PD, Mascola JR. HIV-1 Vaccines Based on Antibody Identification, B Cell Ontogeny, and Epitope Structure. *Immunity* 2018; **48**:855–871.

96 Pegu A, Hessell AJ, Mascola JR, Haigwood NL. Use of broadly neutralizing antibodies for HIV-1 prevention. *Immunol Rev* 2017; **275**:296–312.

97 Kelley RF, O'Connell MP. Thermodynamic Analysis of an Antibody Functional Epitope. *Biochemistry* 1993; **32**:6828–6835.

98 Wu X, Yang Z-Y, Li Y, Hogerkorp C-M, Schief WR, Seaman MS, *et al.* Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science* 2010; **329**:856–61.

99 Scheid JF, Mouquet H, Ueberheide B, Diskin R, Klein F, Oliveira TYK, *et al.* Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding. *Science* 2011; **333**:1633–7.

100 Walker LM, Huber M, Doores KJ, Falkowska E, Pejchal R, Julien J-P, *et al.* Broad neutralization coverage of HIV by multiple highly potent antibodies. *Nature* 2011; **477**:466–70.

101 Mouquet H, Scharf L, Euler Z, Liu Y, Eden C, Scheid JF, *et al.* Complex-type N-glycan recognition by potent broadly neutralizing HIV antibodies. *Proc Natl Acad Sci U S A* 2012; **109**:E3268-77.

102 Walker LM, Phogat SK, Chan-Hui P-Y, Wagner D, Phung P, Goss JL, *et al.* Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. *Science* 2009; **326**:285–289.

103 Huang J, Ofek G, Laub L, Louder MK, Doria-Rose N a, Longo NS, *et al.* Broad and

potent neutralization of HIV-1 by a gp41-specific human antibody. *Nature* 2012; **491**:406–12.

104 Sanders RW, Derking R, Cupo A, Julien JP, Yasmeen A, de Val N, *et al.* 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* 2013; **9**:e1003618.

105 Ward AB, Wilson IA. The HIV-1 envelope glycoprotein structure: nailing down a moving target. Immunol. Rev. 2017; **275**:21–32.

106 Doud MB, Lee JM, Bloom JD. How single mutations affect viral escape from broad and narrow antibodies to H1 influenza hemagglutinin. *Nat Commun* 2018; **9**:1386.

107 Yoon H, Macke J, West AP, Foley B, Bjorkman PJ, Korber B, *et al.* CATNAP: A tool to compile, analyze and tally neutralizing antibody panels. *Nucleic Acids Res* 2015; **43**:W213–W219.

108 Sok D, Pauthner M, Briney B, Lee JH, Saye-Francisco KL, Hsueh J, *et al.* A Prominent Site of Antibody Vulnerability on HIV Envelope Incorporates a Motif Associated with CCR5 Binding and Its Camouflaging Glycans. *Immunity* 2016; **45**:31–45.

109 Sok D, Doores KJ, Briney B, Le KM, Saye-Francisco KL, Ramos A, *et al.* Promiscuous glycan site recognition by antibodies to the high-mannose patch of gp120 broadens neutralization of HIV. *Sci Transl Med* 2014; **6**:236ra63.

110 Garces F, Sok D, Kong L, McBride R, Kim HJ, Saye-Francisco KF, *et al.* Structural evolution of glycan recognition by a family of potent HIV antibodies. *Cell* 2014; **159**:69–79.

111 McLellan JS, Pancera M, Carrico C, Gorman J, Julien J-P, Khayat R, *et al.* Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. *Nature* 2011; **480**:336–343.

112 Wang H, Gristick HB, Scharf L, West AP, Galimidi RP, Seaman MS, *et al.* Asymmetric recognition of HIV-1 envelope trimer by V1V2 loop-targeting antibodies. *Elife* 2017; **6**:1–25.

113 Crooks ET, Osawa K, Tong T, Grimley SL, Dai YD, Whalen RG, *et al.* Effects of partially dismantling the CD4 binding site glycan fence of HIV-1 Envelope glycoprotein trimers on neutralizing antibody induction. *Virology* 2017; **505**:193–209.

114 Lyumkis D, Julien J-P, de Val N, Cupo A, Potter CS, Klasse P-J, *et al.* Cryo-EM Structure of a Fully Glycosylated Soluble Cleaved HIV-1 Envelope Trimer. *Science (80-)* 2013; **342**:1484–1490.

115 Liu Q, Acharya P, Dolan MA, Zhang P, Guzzo C, Lu J, *et al.* Quaternary contact in the initial interaction of CD4 with the HIV-1 envelope trimer. *Nat Struct Mol Biol* 2017; :1–

12.

116 Kim AS, Leaman DP, Zwick MB. Antibody to gp41 MPER Alters Functional Properties of HIV-1 Env without Complete Neutralization. *PLoS Pathog* 2014; **10**:e1004271.

117 Schoofs T, Klein F, Braunschweig M, Kreider EF, Feldmann A, Nogueira L, *et al.* HIV-1 therapy with monoclonal antibody 3BNC117 elicits host immune responses against HIV-1. *Science (80-)* 2016; **352**:997–1001.

118 Wibmer CK, Bhiman JN, Gray ES, Tumba N, Abdool Karim SS, Williamson C, *et al.* Viral escape from HIV-1 neutralizing antibodies drives increased plasma neutralization breadth through sequential recognition of multiple epitopes and immunotypes. *PLoS Pathog* 2013; **9**:e1003738.

119 Mendoza P, Gruell H, Nogueira L, Pai JA, Butler AL, Millard K, *et al.* Combination therapy with anti-HIV-1 antibodies maintains viral suppression. *Nature* 2018; **561**:479–484.

120 Bar-On Y, Gruell H, Schoofs T, Pai JA, Nogueira L, Butler AL, *et al.* Safety and antiviral activity of combination HIV-1 broadly neutralizing antibodies in viremic individuals. *Nat Med* Published Online First: 2018. doi:10.1038/s41591-018-0186-4

121 Liang Y, Guttman M, Williams J a., Verkerke H, Alvarado D, Hu S-L, *et al.* Changes in structure and antigenicity of HIV-1 Env trimers resulting from removal of a conserved CD4 binding site-proximal glycan. *J Virol* 2016; **90**:JVI.01116-16.

122 Townsley S, Li Y, Kozyrev Y, Cleveland B, Hu S-L. Conserved Role of an N-Linked Glycan on the Surface Antigen of Human Immunodeficiency Virus Type 1 Modulating Virus Sensitivity to Broadly Neutralizing Antibodies against the Receptor and Coreceptor Binding Sites. *J Virol* 2016; **90**:829–841.

123 Foley B, Leitner T, Apetrei C, Hahn B, Mizrachi I, Mullins J, *et al.* HIV Sequence Compendium 2017. *Theor Biol Biophys Group, Los Alamos Natl Lab NM* 2017; **LA-UR 17-2**.

124 Verkerke HP, Williams J a., Guttman M, Simonich C a., Liang Y, Filipavicius M, *et al.* Epitope-independent purification of native-like envelope trimers from diverse HIV-1 isolates. *J Virol* 2016; **90**:JVI.01351-16.

125 Boyd DF, Peterson D, Haggarty BS, Jordan APO, Hogan MJ, Goo L, *et al.* Mutations in HIV-1 Envelope That Enhance Entry with the Macaque CD4 Receptor Alter Antibody Recognition by Disrupting Quaternary Interactions within the Trimer. *J Virol* 2015; **89**:894–907.

126 Sethi A, Tian J, Derdeyn CA, Korber B, Gnanakaran S. A Mechanistic Understanding of Allosteric Immune Escape Pathways in the HIV-1 Envelope Glycoprotein. *PLoS Comput Biol* 2013; **9**. doi:10.1371/journal.pcbi.1003046

127 Blish C a., Nguyen MA, Overbaugh J. Enhancing exposure of HIV-1 neutralization epitopes through mutations in gp41. *PLoS Med* 2008; **5**:0090-0103.

128 Back NK, Smit L, Schutten M, Nara PL, Tersmette M, Goudsmit J. Mutations in human immunodeficiency virus type 1 gp41 affect sensitivity to neutralization by gp120 antibodies. *J Virol* 1993; **67**:6897–6902.

129 Bradley T, Trama A, Tumba N, Gray E, Lu X, Madani N, *et al.* Amino Acid Changes in the HIV-1 gp41 Membrane Proximal Region Control Virus Neutralization Sensitivity. *EBioMedicine* 2016; **12**:196–207.

130 O'Rourke SM, Schweighardt B, Phung P, Mesa KA, Vollrath AL, Tatsuno GP, *et al.* Sequences in Glycoprotein gp41, the CD4 Binding Site, and the V2 Domain Regulate Sensitivity and Resistance of HIV-1 to Broadly Neutralizing Antibodies. *J Virol* 2012; **86**:12105–12114.

131 Wagh K, Bhattacharya T, Williamson C, Robles A, Bayne M, Garrity J, *et al.* Optimal Combinations of Broadly Neutralizing Antibodies for Prevention and Treatment of HIV-1 Clade C Infection. *PLoS Pathog* 2016; **12**:1–27.

132 Wagh K, Seaman MS, Zingg M, Fitzsimons T, Barouch DH, Burton DR, *et al.* Potential of conventional & bispecific broadly neutralizing antibodies for prevention of HIV-1 subtype A, C & D infections. *PLoS Pathog* 2018; **14**:1–24.

133 Da Silva J, Coetzer M, Nedellec R, Pastore C, Mosier DE. Fitness epistasis and constraints on adaptation in a human immunodeficiency virus type 1 protein region. *Genetics* 2010; **185**:293–303.

134 Wu NC, Xie J, Zheng T, Nycholat CM, Grande G, Paulson JC, *et al.* Diversity of Functionally Permissive Sequences in the Receptor-Binding Site of Influenza Hemagglutinin. *Cell Host Microbe* 2017; **22**:247–248.

135 Troyer RM, McNevin J, Liu Y, Zhang SC, Krizan RW, Abraha A, *et al.* Variable fitness impact of HIV-1 escape mutations to cytotoxic T lymphocyte (CTL) response. *PLoS Pathog* 2009; **5**:30–34.

136 Dahirel V, Shekhar K, Pereyra F, Miura T, Artyomov M, Talsania S, *et al.* Coordinate linkage of HIV evolution reveals regions of immunological vulnerability. *Proc Natl Acad Sci* 2011; **108**:11530–11535.

137 Adams RM, Kinney JB, Walczak AM, Mora T. Physical epistatic landscape of antibody binding affinity. 2017; :1–15.

138 Otsuka Y, Schmitt K, Quinlan BD, Gardner MR, Alfant B, Reich A, *et al.* Diverse pathways of escape from all well-characterized VRC01-class broadly neutralizing HIV-1 antibodies. *PLOS Pathog* 2018; **14**:e1007238.

139 Barton JP, Goonetilleke N, Butler TC, Walker BD, McMichael AJ, Chakraborty AK.

Relative rate and location of intra-host HIV evolution to evade cellular immunity are predictable. *Nat Commun* 2016; **7**:1–10.

140 Louie RHY, Kaczorowski KJ, Barton JP, Chakraborty AK, McKay MR. Fitness landscape of the human immunodeficiency virus envelope protein that is targeted by antibodies. *Proc Natl Acad Sci* 2018; :201717765.

141 Zanini F, Puller V, Brodin J, Albert J, Neher RA. In vivo mutation rates and the landscape of fitness costs of HIV-1. *Virus Evol* 2017; **3**. doi:10.1093/ve/vex003

142 Shekhar K, Ruberman CF, Ferguson AL, Barton JP, Kardar M, Chakraborty AK. Spin models inferred from patient-derived viral sequence data faithfully describe HIV fitness landscapes. *Phys Rev E Stat Nonlin Soft Matter Phys* 2013; **88**:062705.

143 Gristick HB, von Boehmer L, West Jr AP, Schamber M, Gazumyan A, Golijanin J, *et al.* Natively glycosylated HIV-1 Env structure reveals new mode for antibody recognition of the CD4-binding site. *Nat Struct Mol Biol* Published Online First: 2016. doi:10.1038/nsmb.3291

144 Wagih O. Ggseqlogo: A versatile R package for drawing sequence logos. *Bioinformatics* 2017; **33**:3645–3647.

145 Burton DR, Ahmed R, Barouch DH, Butera ST, Crotty S, Godzik A, *et al.* A Blueprint for HIV Vaccine Discovery. *Cell Host Microbe* 2012; **12**:396–407.

146 Kwong PD, Mascola JR. Human antibodies that neutralize HIV-1: identification, structures, and B cell ontogenies. *Immunity* 2012; **37**:412–25.

147 McCoy LE, Burton DR. Identification and specificity of broadly neutralizing antibodies against HIV. *Immunol Rev* 2017; **275**:11–20.

148 Friedman J, Alam SM, Shen X, Xia SM, Stewart S, Anasti K, *et al.* Isolation of HIV-1-neutralizing mucosal monoclonal antibodies from human colostrum. *PLoS One* 2012; **7**. doi:10.1371/journal.pone.0037648

149 Trkola a, Purtscher M, Muster T, Ballaun C, Buchacher a, Sullivan N, *et al.* Human monoclonal antibody 2G12 defines a distinctive neutralization epitope on the gp120 glycoprotein of human immunodeficiency virus type 1. *J Virol* 1996; **70**:1100–1108.

150 Clotet B, Cooper D. Clinical management of treatment-experienced , HIV- infected patients with the fusion inhibitor enfuvirtide : consensus recommendations. *Aids* Published Online First: 2004. doi:10.1097/01.aids.0000125946.88690.5e

151 Rimsky LT, Shugars DC, Matthews TJ. Determinants of Human Immunodeficiency Virus Type 1 Resistance to gp41-Derived Inhibitory Peptides. *J Virol* 1998; **72**:986–993.

152 Wei X, Decker JM, Liu H, Zhang Z, Arani RB, Kilby JM, *et al.* Emergence of resistant human immunodeficiency virus type 1 in patients receiving fusion inhibitor (T-20)

monotherapy. *Antimicrob Agents Chemother* 2002; **46**:1896–1905.

153 Melby T, Sista P, DeMasi R, Kirkland T, Roberts N, Salgo M, *et al.* Characterization of envelope glycoprotein gp41 genotype and phenotypic susceptibility to enfuvirtide at baseline and on treatment in the phase III clinical trials TORO-1 and TORO-2. *AIDS Res Hum Retroviruses* 2006; **22**:375–385.

154 Reeves JD, Gallo SA, Ahmad N, Miamidian JL, Harvey PE, Sharron M, *et al.* Sensitivity of HIV-1 to entry inhibitors correlates with envelope/coreceptor affinity, receptor density, and fusion kinetics. *Proc Natl Acad Sci* 2002; **99**:16249–16254.

155 Derdeyn CA, Decker JM, Sfakianos JN, Wu X, O'Brien WA, Ratner L, *et al.* Sensitivity of Human Immunodeficiency Virus Type 1 to the Fusion Inhibitor T-20 Is Modulated by Coreceptor Specificity Defined by the V3 Loop of gp120. *J Virol* 2000; **74**:8358–8367.

156 Wensing AM, Calvez V, Günthard HF, Johnson VA, Paredes R, Pillay D, *et al.* 2017 Update of the Drug Resistance Mutations in HIV-1. *Top Antivir Med* 2016; **24**:132–133.

157 Alam SM, Paleos C a, Liao H-X, Scearce R, Robinson J, Haynes BF, *et al.* An inducible HIV type 1 gp41 HR-2 peptide-binding site on HIV type 1 envelope gp120. *AIDS Res Hum Retroviruses* 2004; **20**:836–845.

158 Bonsignori M, Zhou T, Sheng Z, Chen L, Gao F, Joyce MG, *et al.* Maturation Pathway from Germline to Broad HIV-1 Neutralizer of a CD4-Mimic Antibody. *Cell* 2016; **165**:449–463.

159 Korber B, Hraber P, Wagh K, Hahn BH. Polyvalent vaccine approaches to combat HIV-1 diversity. *Immunol Rev* 2017; **275**:230–244.

160 Liao H-X, Lynch R, Zhou T, Gao F, Alam SM, Boyd SD, *et al.* Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. *Nature* 2013; **496**:469–76.

161 Doria-Rose N a, Schramm C a, Gorman J, Moore PL, Bhiman JN, DeKosky BJ, *et al.* Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies. *Nature* 2014; **509**:55–62.

162 Bhiman JN, Anthony C, Doria-Rose NA, Karimanzira O, Schramm CA, Khoza T, *et al.* Supplement - Viral variants that initiate and drive maturation of V1V2-directed HIV-1 broadly neutralizing antibodies. *Nat Med* 2015; **21**:1332–6.

163 Landais E, Murrell B, Briney B, Murrell S, Rantalainen K, Berndsen ZT, *et al.* HIV Envelope Glycoform Heterogeneity and Localized Diversity Govern the Initiation and Maturation of a V2 Apex Broadly Neutralizing Antibody Lineage. *Immunity* 2017; **47**:990–1003.e9.

164 Bhiman JN, Anthony C, Doria-Rose N a, Karimanzira O, Schramm C a, Khoza T, *et al.* Viral variants that initiate and drive maturation of V1V2-directed HIV-1 broadly neutralizing antibodies. *Nat Med* 2015; :1–7.

165 Simonich CA, Doepker L, Ralph D, Williams JA, Dhar A, Yaffe Z, *et al.* Rapid development of an infant-derived HIV-1 broadly neutralizing antibody lineage Cassandra. *bioRxiv* 2018; :0–33.

166 Williams KL, Wang B, Arenz D, Williams JA, Dingens AS, Cortez V, *et al.* Superinfection Drives HIV Neutralizing Antibody Responses from Several B Cell Lineages that Contribute to a Polyclonal Repertoire. *Cell Rep* 2018; **23**:682–691.