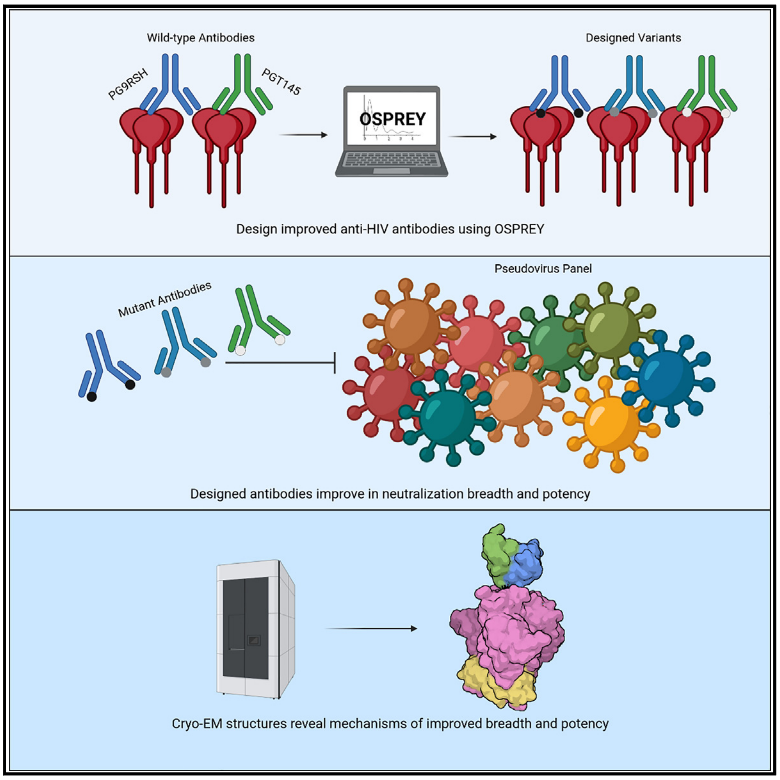
Cell Reports

Article

Improved HIV-1 neutralization breadth and potency of V2-apex antibodies by *in silico* design

Graphical abstract



Design improved anti-HIV antibodies using OSPREY

Designed antibodies improve in neutralization breadth and potency

Cryo-EM structures reveal mechanisms of improved breadth and potency

Highlights

- Antibody variants designed using OSPREY improve neutralization breadth and potency
- Structural and statistical analyses highlight improved side-chain interactions
- Improvements in breadth result from interaction with variable epitope residues

Authors

Graham T. Holt, Jason Gorman,
Siyu Wang, ..., Nicole A. Doria-Rose,
Peter D. Kwong, Bruce R. Donald

Correspondence

pdkwong@nih.gov (P.D.K.),
brd+cell22@cs.duke.edu (B.R.D.)

In brief

Broadly neutralizing antibodies against HIV are promising therapeutics and targets for vaccine elicitation. Using the OSPREY design software, Holt et al. design antibody variants with improved breadth and potency of virus neutralization. They solve bound structures for these variants and provide insight into mechanisms of breadth and potency.

Holt et al., 2023, Cell Reports *42*, 112711
July 25, 2023 © 2023 The Author(s).
https://doi.org/10.1016/j.celrep.2023.112711

Cell Reports

CellPress
OPEN ACCESS

Article

Improved HIV-1 neutralization breadth and potency of V2-apex antibodies by *in silico* design

Graham T. Holt,[1,2,7] Jason Gorman,[3,7] Siyu Wang,[2] Anna U. Lowegard,[1,2] Baoshan Zhang,[3] Tracy Liu,[3] Bob C. Lin,[3] Mark K. Louder,[3] Marcel S. Frenkel,[4] Krisha McKee,[3] Sijy O'Dell,[3] Reda Rawi,[3] Chen-Hsiang Shen,[3] Nicole A. Doria-Rose,[3] Peter D. Kwong,[3,*] and Bruce R. Donald[1,4,5,6,8,*]

[1]Department of Computer Science, Duke University, Durham, NC, USA
[2]Program in Computational Biology & Bioinformatics, Duke University, Durham, NC, USA
[3]Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA
[4]Department of Biochemistry, Duke University, Durham, NC, USA
[5]Department of Mathematics, Duke University, Durham, NC, USA
[6]Department of Chemistry, Duke University, Durham, NC, USA
[7]These authors contributed equally
[8]Lead contact
*Correspondence: pdkwong@nih.gov (P.D.K.), brd+cell22@cs.duke.edu (B.R.D.)
https://doi.org/10.1016/j.celrep.2023.112711

SUMMARY

Broadly neutralizing antibodies (bNAbs) against HIV can reduce viral transmission in humans, but an effective therapeutic will require unusually high breadth and potency of neutralization. We employ the OSPREY computational protein design software to engineer variants of two apex-directed bNAbs, PGT145 and PG9RSH, resulting in increases in potency of over 100-fold against some viruses. The top designed variants improve neutralization breadth from 39% to 54% at clinically relevant concentrations (IC_{80} < 1 μg/mL) and improve median potency (IC_{80}) by up to 4-fold over a cross-clade panel of 208 strains. To investigate the mechanisms of improvement, we determine cryoelectron microscopy structures of each variant in complex with the HIV envelope trimer. Surprisingly, we find the largest increases in breadth to be a result of optimizing side-chain interactions with highly variable epitope residues. These results provide insight into mechanisms of neutralization breadth and inform strategies for antibody design and improvement.

INTRODUCTION

Broad and potent antibodies against HIV-1 show therapeutic promise for preventing viral transmission or infection[1–4] and have been shown to suppress viremia in humans.[5–8] The HIV-1 envelope (Env) apex, comprised of variable loops V1 and V2, is a common target site for anti-HIV-1 broadly neutralizing antibodies (bNAbs)[9–11] despite the high antigen sequence variation at the V1V2 region[12] and the presence of a protective glycan shield.[13,14] These bNAbs form an important category[15,16] that contains the PG9[17–21] and PGT145[22–26] antibody classes, members of which (e.g., PGDM1400[23]) are among the broadest and most potent HIV-1 NAbs thus far identified. Their utility as therapeutics or for prevention, however, would be improved if their potency and breadth were increased. Recent clinical data suggest an *in vitro* IC_{80} of <1 μg/mL for viral strains (measured by TZM-bl assay) to be associated with prevention of transmission in humans.[8]

Extensive structural characterization of bNAb lineages has suggested that breadth of neutralization is conferred by favorable interactions with conserved epitope features. Both the eponymous PG9[17,18] and PGT145[22,24,25] monoclonal antibodies achieve neutralization breadth by targeting conserved structural features on the Env apex.[17,27] PG9 uses its long, axe-like CDRH3 loop to form hydrogen bonds with the C strand of the Env V2 region in a parallel beta-strand conformation and also to interact productively with several apex glycans, including those at Env residues N160, N156, and, in some cases, N173.[17] The beta-strand interaction allows PG9 to maintain favorable contacts with the V2 region despite variation in Env side-chain identities. We previously[28] improved the potency of PG9 by incorporating mutations from the related PG16 antibody to improve interactions with hybrid-type glycans at N173, yielding the antibody PG9-16-RSH (herein, PG9RSH). PGT145 uses its long, needle-like CDRH3 loop to insert sulfated tyrosines into the Env apex hole[24,25] to contact sites of conserved positive charge, both on the C strand and deeper beneath the surface of the Env trimer.[20,24,25] These strategies contrast with those used by members of the VRC38 class of antibodies (a member of the same V1V2 bNAb category), which rely on specific V2 C-strand side-chain interactions to neutralize HIV and, perhaps as a result, exhibit much narrower breadth of neutralization.[29]

The relationship between breadth and potency of neutralization is of considerable interest for antibody design. Although some studies have indicated that improving neutralization against a single antigen can lead to improved neutralization breadth,[1,30–32] other evidence suggests the existence of a trade-off between breadth and potency.[33,34] Exploration of this relationship from a structural perspective is made more challenging by the relative scarcity of high-resolution structure information compared with the extreme antigenic diversity of targets like HIV. One reasonable hypothesis is that concomitant increases in potency and breadth require designing improved interactions with conserved antigen residues.

Methods for improving breadth and potency of antibodies against various targets including HIV-1, dengue, influenza, and severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) have included affinity maturation,[30–35] structure-based *in silico* design,[32,36,37] sequence-based *in silico* design,[38] and hybrid approaches.[1,28,39,40] Previously, we have proposed a knockin humanized mouse model,[41] surface matrix screening,[42] and a structure-based matrix method[43] to yield improved antibodies. *In silico* methods are attractive due to their ability to find improved variants without huge experimental effort and have found success in improving both breadth and potency of neutralization. Researchers have improved the breath and potency of cross-reactive anti-dengue,[38] cross-reactive anti-influenza,[37] and anti-HIV-1[32] antibodies using computational strategies. Hybrid methods that include *in silico* design have also achieved remarkable improvements.[1]

In this work, we design the PG9RSH and PGT145 anti-HIV-1 bNAbs for improved potency and breadth using our computational protein design software OSPREY.[44] We use predicted potency for the BG505 strain as a proxy for predicted neutralization breadth, and we computationally optimize interactions with both conserved and non-conserved epitope residues. We present and characterize three bNAb single-mutation variants, compare them with both wild-type (PG9RSH and PGT145) and best-in-class (PGDM1400) antibodies, and observe measured improvements in breadth or potency relative to wild type. One of these three point mutations has been previously characterized in PG9 but not in PG9RSH.[32] We determine cryoelectron microscopy (cryo-EM) structures for these three designed variants to provide atomic-level insight into increases in breadth and potency. The largest improvements in median potency ($\approx$ 3-fold IC_{50}, $\approx$4-fold IC_{80}) occurred for PG9RSH variant DU025, which achieves neutralization breadth and potency rivaling that of the best-in-class antibody PGDM1400. Surprisingly, we find that the largest improvements in breadth occurred for a variant that optimizes interactions with variable epitope residues. This variant, PGT145 DU303, lost subtype potency for clade B but nonetheless improved overall breadth of neutralization from 39% (wild type) to 54% at clinically relevant concentrations (IC_{80} < 1 μg/mL). For this designed antibody, we observed increases in potency of >100-fold for six pseudoviruses across five clades. Moreover, the median improvement in IC_{80} (across 208 strains) was over 3-fold.

RESULTS

OSPREY predicts mutations with improved affinity for PGT145 and PG9RSH

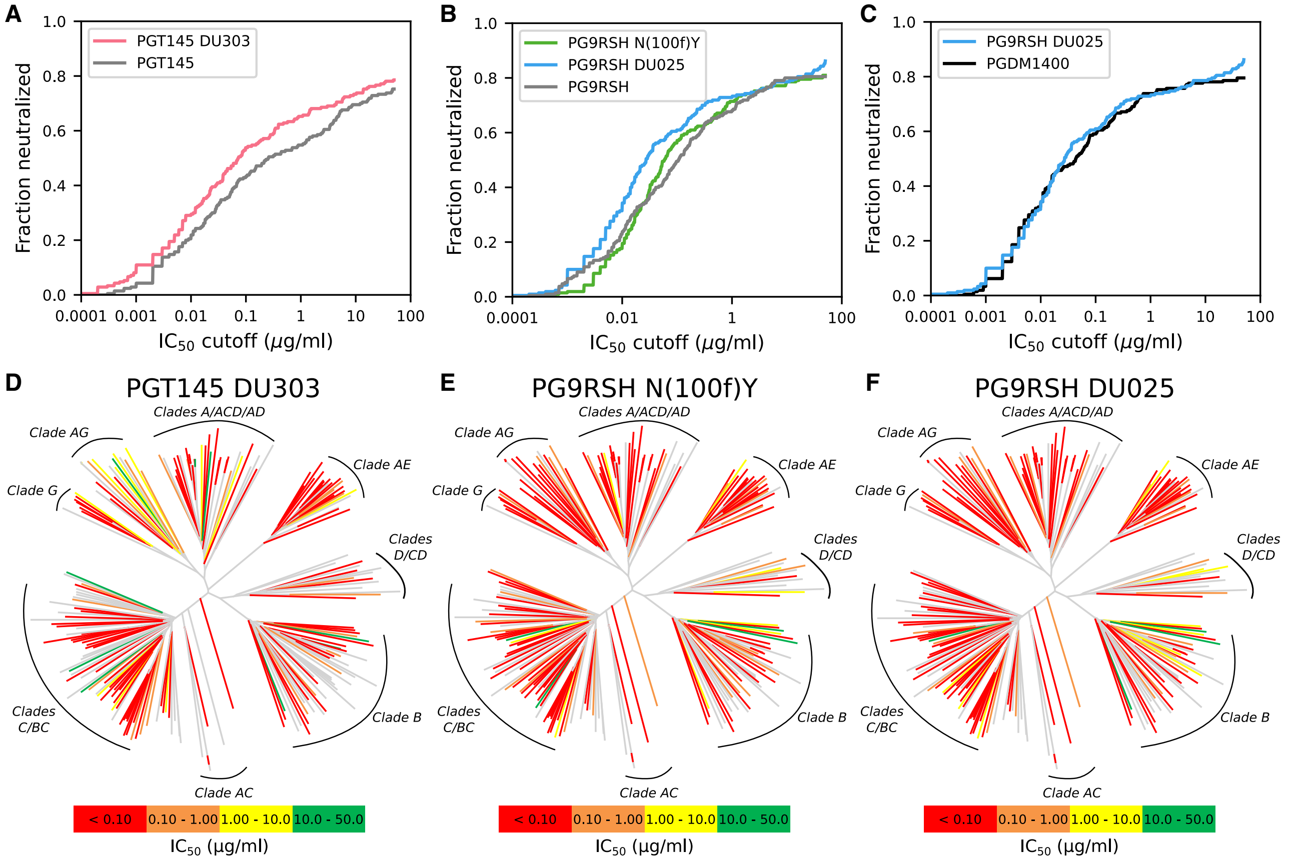
We used OSPREY to design variants of the PGT145 and PG9 antibodies using structures of each antibody bound to the BG505 SOSIP Env trimer (PDB: 5U1F, 5VJ6). We predicted the effect of antibody mutations on antigen-binding affinity by computing K^* scores[45–48] for both wild-type and mutant antibodies at 9 positions (Table S1), pruning or computing scores for approximately 14,000 sequences. An increase in score relative to wild type predicts an increase in K_a , i.e., improved binding. In general, antibody binding affinity for the Env spike correlates well with neutralization potency.[49] These designs predicted that PGT145 N(100l)D (variant DU303), PG9RSH N(100f)Y, and PG9RSH Y(100k)D (variant DU025), among other substitutions, would improve the neutralization potency of PGT145 and PG9RSH.

Double-mutation designs of PGT145 at residues F(100d) and N(100l) predicted that negatively charged substitutions at position 100l would improve binding affinity (Figure S1; Table S2). To avoid destabilization of the antibody CDRH3 conformation, sequences for which the lower bound on the partition function for the antibody state (Z_{Ab}) was less than $10^{17.7}$ were excluded from consideration. Mutations N(100l)D, F(100d)H/N(100l)D, and N(100l)E were all predicted to increase the K^* score. Notably, these substitutions place a negatively charged side chain at residue 100l, perhaps leveraging proximity to the (generally) positively charged Env residues 166 and 169 (see Figure 3C). Additional designs of PGT145 are discussed in Document S2.

A single-residue design of PG9RSH at residue N(100f) predicted that substitutions to Trp, Met, Tyr, His, or Phe would improve binding affinity (Figure S2A; Table S3). The mutation N(100f)Y was discovered previously and was shown to improve breadth and potency of neutralization in PG9.[50] An additional single-residue design of PG9RSH at residue Y(100k) predicted that large or negatively charged substitutions (Trp, Asp, or Glu) would improve binding affinity (Figure S2B; Table S4). Additional designs of PG9RSH are discussed in Document S2.

Neutralization assessment reveals improvements in breadth and potency

Based on these designs, we selected 10 and 34 variants of antibodies PGT145 and PG9RSH, respectively, for small-panel neutralization assays (Tables S5 and S6; Data S1). These variants were selected by prioritizing variants with high K^* scores and (to a lesser extent) high unbound-state partition functions,[51] and in some cases, promising mutations from different designs were combined. We then selected variants DU303 (PGT145 N(100l)D), PG9RSH N(100f)Y, and DU025 (PG9RSH Y(100k)D) for assay against a large panel of 208 pseudoviruses to further characterize their potency and breadth of neutralization (Data S2). These variants were selected based on the number of pseudoviruses neutralized with an $IC_{80} \leq 50$ μg/mL, the median IC_{80} value, the number of pseudoviruses neutralized with an $IC_{50} \leq$ 50 μg/mL, and the median IC_{50} value (listed in order of importance). For example, variants N(100f)Y and DU017 were selected over DU014 because DU014 performs comparatively poorly as measured by IC_{80}. Although its IC_{50} neutralization breadth appears to be greater, the additional neutralized virus is neutralized relatively poorly (IC_{50} = 36.1 μg/mL) and disappears when measured using IC_{80}. Furthermore, the median IC_{50} and IC_{80} values of DU014 are greater than for N(100f)Y (less potent). Additionally, the sequence and neutralization diversity of the set of



G

	N(100f)Y	DU025	PG9RSH	DU303	PGT145	PGDM1400
Breadth IC$_{80}$ < 1.0 μg/mL	62%	63%	53%	54%	39%	63%
Median IC$_{80}$ (μg/mL)[1]	0.086	0.058	0.227	0.090	0.276	0.047
Mean IC$_{80}$ (μg/mL)[2]	0.106	0.078	0.236	0.116	0.343	0.069

[1] Median calculated for samples with IC$_{80}$ < 50 μg/mL.
[2] Geometric mean calculated for samples with IC$_{80}$ < 50 μg/mL.

Figure 1. Large-panel neutralization breadth and potency for PGT145 and PG9RSH variants

Neutralization breadth and potency of PGT145 and PG9RSH variants assayed on a panel of 208 pseudoviruses.

(A and B) Breadth/potency curves for PGT145 and PG9RSH variants and controls, respectively. Curves represent the fraction of pseudoviruses that were neutralized with IC$_{50}$ smaller than the given cutoff. An increase in breadth and potency is indicated by a shift upward and left. PGT145 variant DU303 and PG9RSH variants N(100f)Y and DU025 improve breadth and potency relative to wild type. For IC$_{80}$ curves, see Figure S3.

(C) Breadth/potency curves for PG9RSH DU025 and PGDM1400. Despite slightly weaker median and mean neutralization potency, the overall breadth and potency of DU025 rival that of the best-in-class PGDM1400 antibody.

(D–F) Neutralization dendrograms for variants DU303, PG9RSH N(100f)Y, and DU025, and respectively. Pseudoviruses are grouped into clades by sequence similarity, forming a tree graph. Internal branches in the tree, which denote groups of viruses, are colored gray. Terminal branches, corresponding to a single pseudovirus, are colored by IC$_{50}$, where high neutralization potency is indicated by warm colors, low potency is indicated by cool colors, and lack of neutralization is indicated by gray.

(G) Summary of large-panel neutralization breadth and potency for variants and controls, measured by IC$_{80}$.

variants to be characterized was considered. DU303, DU025, and PG9RSH N(100f)Y improved over wild-type activity in both breadth and potency of neutralization, although PG9RSH N(100f)Y achieved only small increases in breadth.

DU303 neutralized more pseudovirus strains with higher potency (Figure 1A). DU303 increased neutralization potency against the BG505 strain by 3- and 25-fold as measured by IC$_{50}$ and IC$_{80}$, respectively: IC$_{50}$ decreased from 0.010 to 0.003 μg/mL, and IC$_{80}$ decreased from 0.253 to 0.010 μg/mL (Data S2). Median neutralization potency across the large panel increased by 2- and 3-fold (IC$_{50}$ and IC$_{80}$): median IC$_{50}$ decreased from 0.053 to 0.024 μg/mL, and median IC$_{80}$ decreased from 0.276 to 0.090 μg/mL (Tables S7 and S8). DU303 improved neutralization breadth relative to PGT145: the percentage of tested pseudoviruses with measurable neutralization (IC$_{50}$ < 50μg/mL) increased from 75% to 79% (Table S7). Interestingly, the improvement in breadth relative to PGT145 was more pronounced when evaluated at a cutoff with clinical relevance[8]: the percentage of viruses neutralized with IC$_{80}$ < 1 μg/mL increased from 39% to 54% (Table S8).

PG9RSH N(100f)Y increased median potency of neutralization but only slightly increased breadth (Figure 1B). PG9RSH N(100f)

Y increased neutralization potency against the BG505 strain by 2.8-fold as measured by IC_{80} but showed no appreciable change in IC_{50}: IC_{80} decreased from 0.065 to 0.023 μg/mL, and IC_{50} remained at 0.009 μg/mL (Data S2). Median neutralization potency across the large panel increased by 1.9- and 2.6-fold (IC_{50} and IC_{80}): median IC_{50} decreased from 0.047 to 0.025 μg/mL, and median IC_{80} decreased from 0.227 to 0.086 μg/mL (Tables S7 and S8). PG9RSH N(100f)Y slightly improved neutralization breadth relative to PG9RSH: the percentage of tested pseudoviruses with measurable neutralization ($IC_{50} < 50$ μg/mL) increased from 81% to 83% (Table S7). However, the improvement in breadth relative to PG9RSH was larger when evaluated for pseudoviruses neutralized with an $IC_{80} < 1$ μg/mL: breadth increased from 53% to 62% (Table S8).

DU025 increased potency and breadth of neutralization (Figure 1B), and the resulting breadth-potency plot is qualitatively similar to that for the best-in-class PGDM1400 antibody (Figure 1C). DU025 increased neutralization potency against the BG505 strain by 2.2- and 6.5-fold as measured by IC_{50} and IC_{80}, respectively: IC_{50} decreased from 0.009 to 0.004 μg/mL, and IC_{80} decreased from 0.065 to 0.010 μg/mL. Median neutralization potency across the large panel increased by 2.7- and 3.9-fold (IC_{50} and IC_{80}): median IC_{50} decreased from 0.047 to 0.017 μg/mL, and median IC_{80} decreased from 0.227 to 0.058 μg/mL (Tables S7 and S8). DU025 also improved neutralization breadth relative to PG9RSH: the percentage of tested pseudoviruses with measurable neutralization ($IC_{50} < 50$ μg/mL) increased from 81% to 87% (Table S7). Again, the improvement in breadth relative to PG9RSH was larger when evaluated at the clinically relevant cutoff of $IC_{80} < 1$ μg/mL: breadth increased from 53% to 63% (Table S8). While the mean and median potencies for DU025 remained slightly weaker than for PGDM1400, overall, the breadth and potency of DU025 rivaled that of best-in-class antibodies.

Strains with positively charged side chains at Env residue 169 are neutralized more potently by DU303

To further characterize the changes in neutralization activity for variant antibodies relative to PGT145 and PG9RSH, we computed the fold decrease in IC_{50} for each tested pseudovirus (Figure S4). We observed increases in potency of >100-fold for seven virus-antibody pairs across five clades, including 988-, 643-, and 228-fold improvements in IC_{50} for strains CAP256.206.C9, 16936-2.21, and CH038.12, respectively (Table S9). Analysis revealed a wide distribution of improvements in neutralization for DU303, with populations of strains with either slightly reduced or greatly improved neutralization. Conversely, both PG9RSH N(100f)Y and DU025 were characterized by a narrow distribution of fold decrease in IC_{50} over tested pseudoviruses. The largest fold decrease in IC_{50} for DU303 was 988-fold, occurring for the clade C pseudovirus CAP256.206.C9. The largest improvements in neutralization for PG9RSH N(100f)Y and DU025 were 63- and 110-fold, respectively, both occurring for the clade C pseudovirus ZM233.6. Examination of the fold change in neutralization by clade (Figures S5A, S6, and S7C) revealed the improvements in breadth and potency for DU303 to be non-uniform: marked improvements are evident for clades C/BC, A/AD, and D/CD, but decreases in neutralization occurred for clade B. On the other hand, the same analysis for both PG9RSH N(100f)Y (Figures S5B, S6, and S7A) and DU025 (Figures S5C, S6, and S7B) revealed most clades to show slight increases in neutralization with no clear pattern of decreased neutralization.

A gradient-boosted trees classifier predicted the changes in neutralization of pseudoviruses between wild-type PGT145 and DU303 and indicated these changes to be associated with the amino acid identity at Env residue 169. To investigate Env sequence features that may explain changes in neutralization for DU303, PG9RSH N(100f)Y, and DU025 relative to their ancestors, we trained gradient-boosting tree models (similar to the approach in Rawi et al.[52]) to predict the sign of the change in neutralization for each variant based on pseudovirus Env sequences. Models were evaluated using repeated 10-fold nested cross-validation (Table S10; Figure S8). The model for DU303 performed well, with a mean area under the curve (AUC) of 0.807 ± 0.128, but models for PG9RSH N(100f)Y and DU025 performed poorly, with mean AUCs of 0.521 ± 0.134 and 0.571 ± 0.134, respectively. To identify residues on Env that were important for the improved neutralization of DU303, we evaluated the permutation importance (PI)[53] of Env residues. K169 and R169 were the important ($PI > 0.05$) features that were associated with an increase in neutralization for DU303 relative to PGT145 (Figure S9). Separating the large-panel data by residue identity at Env residue 169 revealed two distinct populations of viruses (Figure S10): Viruses with a positively charged residue at Env residue 169 were more potently neutralized by DU303 than by PGT145, while viruses without a positively charged residue at this position were less potently neutralized by DU303. We propose a structural mechanism for this observation in Document S2 supported by homology modeling.

Cryo-EM structures of BG505 DS-SOSIP.664 bound by DU303, PG9RSH N(100f)Y, and DU025 reveal improved side-chain interactions

We solved cryo-EM structures of PGT145 variant DU303, PG9RSH N(100f)Y, and PG9RSH variant DU025 in complex with the BG505 DS-SOSIP.664 Env trimer (Figure 2). Three-dimensional reconstructions yielded resolutions of 3.58, 3.40, and 3.75 Å, respectively (Table S11), and local resolutions ranged between 3 and 9.3, 2.7 and 6.3, and 3.18 and 7.5 Å, respectively (Figures S11–S13). The trimer apex and antibody CDRH3 loop were well resolved in all cases and indicated binding modes consistent with previous structures of PGT145 and PG9RSH (Figure 2). These structures revealed details for key interactions between the HIV Env apex and the DU303, PG9RSH N(100f)Y, and DU025 variant antibodies. An extended structural analysis can be found in Document S2.

DU303 improves side-chain interactions with HIV Env residues 166 and 169 by introducing the N(100l)D mutation. Cryo-EM maps show well-resolved electron density for gp120 residues R166 and K169 but reveal ambiguity in the precise side-chain placements of residues D(100l) and F(100d) (Figure S14). The atomic model of DU303 indicates that D(100l) could form electrostatic interactions with gp120 residues R166 and K169: the side-chain nitrogen of K169 lies 5.1 Å from a side-chain carboxyl oxygen of D(100l) (Figure 3D). Similarly, one of the side-chain

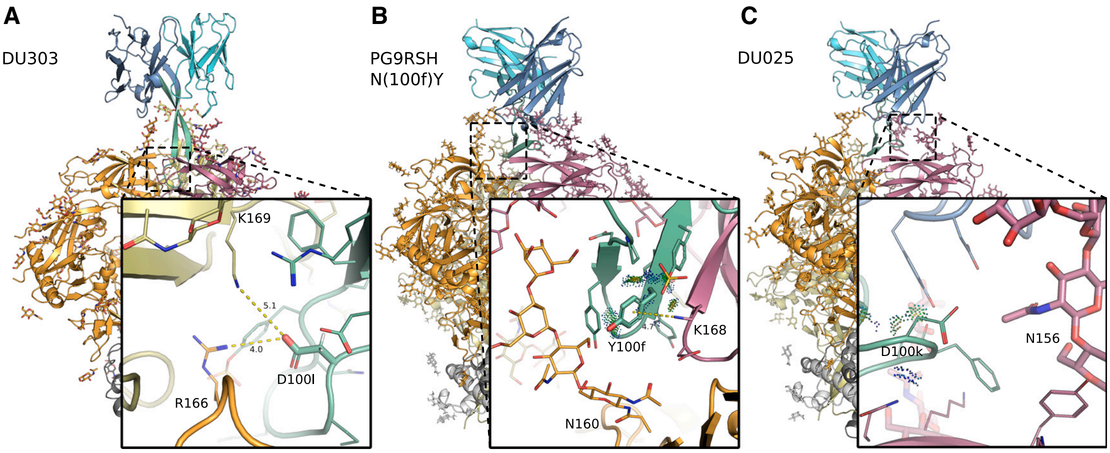


Figure 2. Cryo-EM structures of PG9RSH and PGT145 variants in complex with BG505 DS-SOSIP Env trimer

Backbone shown in ribbon representation with glycans, and amino acids shown as sticks or lines. Env subunits colored with warm colors and grays, and antibodies shown in cool colors. CDRH3 (residues 95–102) is shown in green. Distances (Å) are shown with dotted yellow lines, and energetic interactions are shown with Probe dots. Members of the PG9RSH (A and B) and PGT145 lineages (C) interact with the trimer apex and are characterized by a negatively charged CDRH3 that is hammer-like or extended, respectively.

(A) PGT145 mutation N(100l)D forms more favorable electrostatic interactions with gp120.

(B) The PG9RSH N(100f)Y mutation creates interactions with a side-chain nitrogen from gp120 residue K168, forming geometry consistent with a π-cation interaction.

(C) The PG9RSH Y(100k)D mutation forms long-range interactions with polar and positively charged residues Q170 and K305.

nitrogens of R166 lies 4 Å from a side-chain carboxyl oxygen of D(100l). The position of these side chains suggests that the negatively charged D(100l) forms favorable interactions with positively charged residues on gp120 to improve breadth and potency of neutralization.

PG9RSH N(100f)Y improves side-chain interactions with Env residue 168. Electron density maps show well-resolved density for gp120 residues D167, K168, and K169, along with the first two N-acetylglucosamine (GlcNAc) sugars of gp120 glycan N160 (Figure S15A). Density corresponding to bNAb residues is more ambiguous: peaks between the modeled side-chain locations of residues Y(100f) and Y(100a) suggest the presence of alternate rotamer configurations. Examination of low-density peaks (0.5 σ) reveals a small peak in density of the second GlcNAc of glycan N160 near the modeled location of Y(100f), suggesting interactions between Y(100f) and the glycan shield (Figure S15B). The atomic model of PG9RSH N(100f)Y fit to the density map indicates that the primary interaction between Y(100f) and gp120 is a π-cation interaction with residue K168 (Figure 3E). The ammonium nitrogen of K168 lies 4.7 Å from the center of the Y(100f) π system, and the angle between the distance vector and the ring normal vector is approximately 20°, which is representative of typical π-cation geometry.[54] Y(100f) also forms van der Waals interactions with antibody residues P99, Y(100a), and TYS(100h). The side-chain geometry suggests that the aromatic Y(100f) side chain participates in a π-cation interaction with the positively charged K168 to improve potency of neutralization.

DU025 may improve long-range side-chain interactions or glycan interactions by introducing the Y(100k)D mutation. The electron density around the side chains of D(100k), Q170, K305, and Y173 are well resolved, along with the core of glycan N156 (Figure S16A). Interestingly, three unassigned density peaks arise in the groove between the V2 and V3 loops at both 1.2 and 3 σ, which could indicate the presence of solvent at this interface (Figure S16B). Furthermore, a bridge of density at 1 σ arises between the modeled locations of Env residues Q170 and R308, hinting at long-range or solvent-mediated interactions. These data suggest that residue D(100k) may form long-range or solvent-mediated interactions with residues Q170 and K305, which lie at distances of 4.1 and 6.8 Å, respectively, in the atomic model.

OSPREY predictions are validated by cryo-EM structures

OSPREY designs of antibody variants correctly predicted side-chain interactions. For DU303, predicted interaction distances between D(100l) and K169 and R169 in the OSPREY low-energy ensemble (LEE) differed by at most one angstrom from distances in the experimental model (Figures 3A and 3D). The side-chain orientations of this system were qualitatively similar between the LEE and the experimental model, indicating that OSPREY correctly predicted the structural consequences of the N(100l)D substitution. For PG9RSH N(100f)Y, side-chain locations in the LEE were qualitatively similar to those in the experimental model (Figures 3B and 3E). Interestingly, multiple rotamers of Y(100f) and Y(100a) appeared in the LEE, resulting in a conformation in which these side chains have rotated and stacked. The overall correspondence between the LEE and the experimental model indicated that OSPREY correctly predicted the structural consequences of the N(100f)Y substitution. Designs of DU025 predicted interactions with Q170 and R308, but differences in loop backbone conformation resulted in a change of environment near residue 100k. As a result, although the design

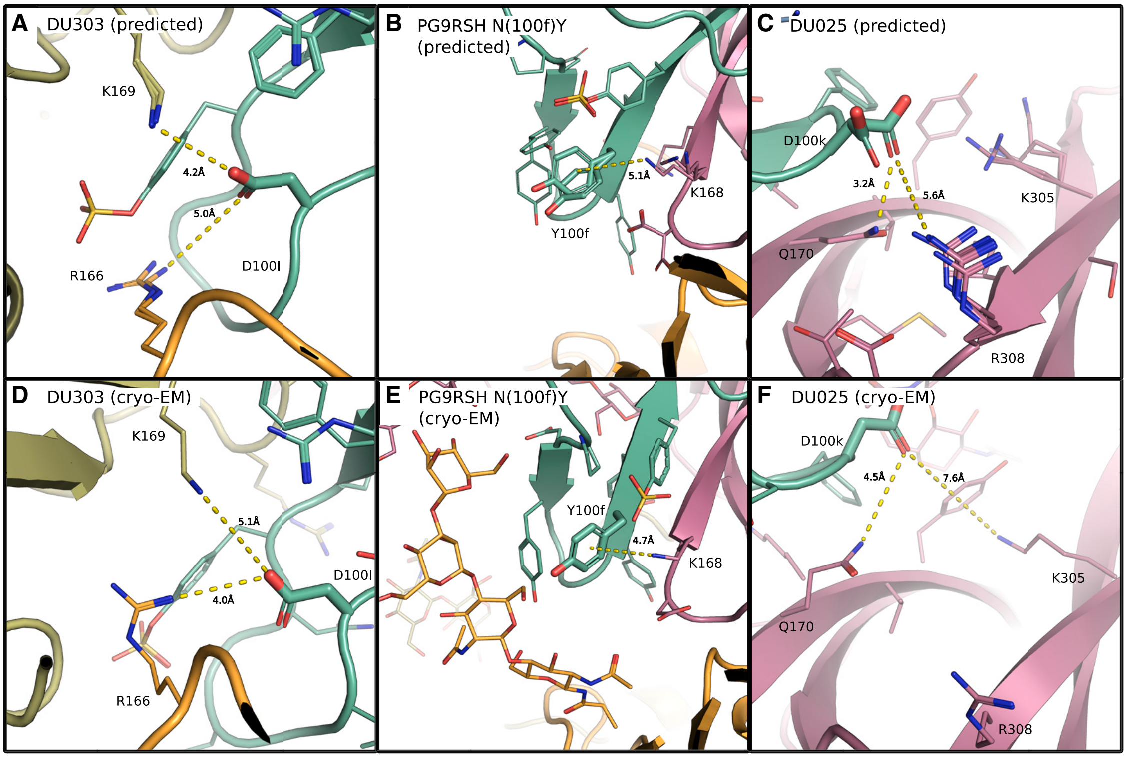


Figure 3. OSPREY design ensembles correctly predicted structural features for PG9RSH and PGT145 variants

Ten members of the low-energy ensemble (LEE) predicted by OSPREY are shown for variants of PGT145 (A) and PG9RSH (B and C) above corresponding cryo-EM structures (D–F). Backbones are shown as ribbons with amino acids shown as lines or as sticks with Env subunits colored with warm colors, and antibody CDRH3 loops (residues 95–102) are shown in green. Distances (Å) are shown with dotted yellow lines.

(A) PGT145 mutation N(100l)D is predicted to form electrostatic interactions with gp120 residues R166 and K169. A carboxyl oxygen of D(100l) lies 5 and 4.2 Å from side-chain nitrogens of gp120 residues R166 and K169, respectively. Despite a lateral translation of the CDRH3 loop relative to the trimer apex, the LEE correctly predicts features of the experimental structure (F).

(B) The PG9RSH N(100f)Y mutation creates interactions with gp120 residue K168. The side-chain amino nitrogen of K168 lies 5.1 Å from the ring plane of Y(100f), forming geometry consistent with a π-cation interaction. The LEE correctly predicts interactions found in the experimentally determined structure (D).

(C) The PG9RSH Y(100k)D mutation is predicted to form electrostatic interactions with polar and charged residues on gp120. A carboxyl oxygen of D(100k) lies 3.2 Å from the side-chain nitrogen of Q170 and 5.6 Å from R308. The LEE correctly predicts interactions with Q170, but a translation and rotation of the CDRH3 loop places R308 further away (E).

ensemble correctly predicted that D(100k) creates long-range electrostatic interactions and correctly predicted a favorable interaction with Q170 (Figures 3C and 3F), the predicted interaction between D(100k) and R308 is not supported by the cryo-EM structure. Instead, shifts in the backbone create interactions between D(100k) and K305. However, the overall quality and type of interactions formed by D(100k) in the design ensemble were in fact consistent with the experimental structure. For an extended analysis of structural correspondence between the LEE and cryo-EM structures, see Document S2.

DISCUSSION

In this work, we tested our ability to redesign apex-directed anti-HIV bNAbs for improved neutralization breadth using the OSPREY protein design software. Ideally, to model structural variation, we would incorporate high-resolution co-crystal structures of bNAb:Env complexes from all tested strains into the design process, but such structures are scarce and difficult to obtain. In fact, at the time of design, the BG505 strain was the only HIV genotype represented in experimental structures of PGT145 and PG9 complexed with trimeric Env. Additionally, complex structures for poorly neutralized strains may be difficult to solve due to affinity issues. Previous studies indicated that improving neutralization potency against a single virus strain, either by affinity maturation[30] or by design,[1,31,32] can result in improved neutralization breadth. Therefore, we used the predicted affinity for the BG505 DS-SOSIP.664 trimer as a proxy for neutralization breadth during the design process. Assessment on a panel of 208 Env pseudoviruses indicated that three designed variants exhibited improved neutralization breadth and potency. We solved structures of these three variants bound to the BG505 DS-SOSIP.664 trimer to investigate the mechanisms of improved neutralization potency. We additionally investigated relationships between Env epitope residue characteristics and neutralization potency to draw conclusions about mechanisms of improved breadth. Surprisingly, mutations that optimized interactions with variable epitope residues resulted in the largest improvements in breadth.

Single mutations improve antibody neutralization of BG505

Experimental characterization of bNAb variants showed that DU303, PG9RSH N(100f)Y, and DU025 improved or maintained neutralization potency for the BG505 pseudovirus, with the most notable improvements observed by IC_{80} measurements. Cryo-EM structures of each variant bound to the BG505 trimer indicate that OSPREY designs improved side-chain interactions. The PGT145 N(100l)D mutation (DU303) improved electrostatic interactions with the Env apex residues R166 and K169, improving charge complementarity. The PG9RSH N(100f)Y substitution created a π-cation interaction with Env residue K168 and may also interact with glycan N160. Improved stability of the CDRH3 loop has also been suggested for this mutation.[32] Finally, the PG9RSH Y(100k)D mutation (DU025) improved side-chain interactions with the polar Env residue Q170 and glycan N156. The interface around residue D(100k) is difficult to resolve, which may be indicative of a mobile or solvent-accessible environment. Unassigned electron-density peaks suggest that D(100k) may also form solvent-mediated interactions with K305 and perhaps even R308. All three designs were successful in improving neutralization potency against BG505 by optimizing side-chain interactions. The general correspondence between OSPREY-generated design ensembles and cryo-EM structures also indicated that our algorithms accurately modeled both environment and side-chain interactions at the PG9RSH and PGT145 epitopes. This is interesting given the low resolution of the design input structures and may be due to the fact that our algorithms are more sensitive to the input backbone conformation than the side chains, which are more difficult to resolve experimentally.

bNAb variants show different patterns of improvement in neutralization breadth

The three best antibody variants improved neutralization breadth across a panel of 208 HIV pseudoviruses but differed in the pattern and extent of change in breadth. DU303 improved neutralization for most clades but sacrificed some subtype potency for clade B. PG9RSH N(100f)Y and DU025, on the other hand, increased neutralization in a relatively uniform manner across all clades. Interestingly, DU303 and DU025 improved overall breadth to a greater extent despite relatively low conservation of their Env epitope residues. Conversely, PG9RSH N(100f)Y resulted in smaller improvements despite the high conservation of Env residues that interact with the mutated antibody residue 100f (Figure S17).

PGT145 DU303 improved breadth by improving potency against "sensitive" strains containing a lysine or arginine at residue 169 while slightly decreasing potency against "resistant" strains with different substitutions at this epitope residue. Sensitive strains were more potently neutralized by ≈5-fold (geometric mean), while resistant strains were less potently neutralized by ≈2-fold (Figure S10). Because the effect of improving neutralization against sensitive strains was larger than the effect of decreasing neutralization against resistant strains, and because there were more sensitive strains than resistant strains, the aggregate effect of the N(100l)D substitution was a gain of breadth and potency. These observations explain the loss of subtype potency for clade B, in which Env residue 169 is predominantly hydrophobic (valine, methionine) and unlikely to interact favorably with D(100l).

PG9RSH DU025 improved breadth of neutralization by improving interactions with variable residues on the Env V2 and V3 loops and with the conserved glycan N156. Analysis across a large panel of pseudoviruses revealed no major decreases in subtype potency, despite the relative variability of the epitope residues in proximity to residue D(100k) (Figure S17). This could indicate that the Y(100k)D mutation improves breadth by improving interactions with the conserved glycan N156 or by improving interactions with the variable Env residues 170 and 305 in a manner that is tolerant of variation. Most panel strains have polar or charged residues at Env positions 170 (Q, K, or R) and 305 (K, R, or T), and it is possible that D(100k) interacts favorably with any of these amino acids, especially if interactions were to be solvent mediated.

PG9RSH N(100f)Y slightly improved breadth of neutralization by improving interactions with Env residue 168 and glycan N160. Overall, the slight improvement in breadth did not appear to sacrifice subtype potency, likely because the N(100f)Y substitution interacts with highly conserved Env features. The improvement in breadth was small relative to DU303 or DU025, suggesting that these interactions may already be highly optimized in parent PG9RSH.

Improvements in breadth did not require residue conservation

These examples confirm previous observations[1,30–32] that antibody neutralization breadth can be increased by improving potency for a single "design antigen." One intuitive explanation for this phenomenon is that the design antigen contains residues that are conserved across the entire antigen population. However, for our designs, epitope residue conservation did not appear to be critical for improving breadth. At least one variant (DU303) optimized interactions with epitope residues that were among the least conserved across the 208-strain test panel. Another (DU025) interacts with an epitope containing a conserved glycan, but structures suggest interactions with multiple non-conserved residues.

There are several possible explanations for increased breadth without epitope-residue conservation. First, the energetic benefit for strains with residues that match the BG505 sequence could be larger than the energetic cost for strains that do not match. For example, for DU303, strains containing Env K169 (like BG505) have improved electrostatic interactions with antibody residue D(100k). Although strains with Env V169 would not realize the same energetic benefit, they would be unlikely to suffer from large steric clashes because of the smaller size of valine relative to lysine. As a result, optimizing for design antigen residues would improve overall breadth, even though they are non-conserved. If this is correct, we would expect improvements in breadth to be less likely to occur when the design antigen contains small epitope residues at a design site.

Second, the presence of long-range, possibly solvent-mediated interactions in these designs could allow for favorable energetic interactions with multiple amino acids at epitope sites. As a result, the local chemical environment around design residues could be more conserved than the epitope amino acid identity. In this case, we would expect improvements in breadth to be less likely to occur when optimizing close-range interactions with the design antigen.

Finally, it is possible that natural selection has already optimized interactions with highly conserved residues. The difference between Env residue identity distributions for donor virus populations and test panels depends in part on residue conservation. If natural selection is effectively optimizing antibody residues for donor virus populations, we would expect improvements in test panel neutralization breadth to be larger or more likely when optimizing interactions with non-conserved epitope residues. Further research will be necessary to evaluate these hypotheses.

Comparison to previous work

Various other methods for computational antibody design have been proposed and implemented. Hybrid knowledge-based and computational-structure-based design methods like RAbD,[55] OptCDR[56] and OptMAVEn,[57,58] and abDesign[59] graft experimentally determined CDR loop backbone structures, then use computational structure-based design to optimize loop sequences. However, these approaches are likely most useful for *de novo* antibody design, and the variability of antibody CDRH3 loops makes these approaches suboptimal for antibodies that rely heavily on CDRH3 contacts for neutralization[55] such as PG9RSH or PGT145. Deep learning approaches[60] have been shown to model loops more effectively than some grafting methods. Straightforward applications of general protein design frameworks or algorithms like Rosetta [32] or variants of dead-end elimination or A*[36] have also been used to design antibodies in a more conservative manner. In some cases, modifications are necessary: Tidor and co-workers found that relying on electrostatic interactions improved the performance of their algorithms.[36] Our approach—a straightforward application of OSPREY—is most similar to these latter works, although unlike the designs using Rosetta,[32] our algorithms provably return optimal sequences and structures with respect to the input model. Finally, some approaches explicitly use multistate design to optimize for neutralization across a wide panel of antigens.[37,61] However, this requires either a varied set of experimentally determined antibody:antigen structures or extensive homology modeling, which can degrade performance by inaccurately modeling important structural interactions. Although this is a promising future direction, our results and others[30,32] indicate that improved neutralization breadth can also be achieved through design for potency against one representative antigen.

Future directions

Until more structures of bNAbs in complex with HIV Env trimers from diverse strains become available, methods for structure-based design for breadth will likely continue to rely on information gleaned from one or a few representative antigens. Can the method presented herein be applied to any antigen genotype or site to successfully improve breadth? If not, how can we differentiate between genotypes or sites that are more “canonical” (i.e., for which design for potency can result in increased breadth) versus those that are less? We have suggested several hypotheses based on our new data, but additional work will be required to evaluate these hypotheses. Orthogonally, our method could potentially be used not only for therapeutic antibody improvement but also to assess the projected importance of rare mutations in antibodies targeted for elicitation through vaccination or, alternatively, to identify important immunogen variants.

Limitations of the study

As mentioned above, due to the scarcity of trimeric structures, these designs relied exclusively on structures of antibodies bound to the BG505 SOSIP trimer. Additionally, like most other structures of bNAbs in complex with trimeric Env, our solved structures of improved antibody variants also include the BG505 SOSIP trimer. Although we attempt to mitigate these limitations by characterizing neutralization over a large set of strains and by homology modeling approaches, our understanding of the mechanisms of neutralization breadth for these antibodies and others would be improved by complex structures that span a more diverse set of stable Env trimers.[62] Finally, neutralization potency and breadth are not themselves sufficient for therapeutic efficacy—other properties, including thermostability and lack of autoreactivity, are critical.[1] Further characterization of the antibodies developed herein will be needed to assess their therapeutic potential.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Redesign of PG9RSH and PGT145
  - Antibody variant expression and purification
  - Pseudovirus neutralization assays
  - Cryo-EM data collection, structure determination, and refinement
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Predicting change in neutralization from Env sequence
  - Visualization and figure generation

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2023.112711.

ACKNOWLEDGMENTS

We thank J. Baalwa, D. Ellenberger, F. Gao, B. Hahn, K. Hong, J. Kim, F. McCutchan, D. Montefiori, L. Morris, E. Sanders-Buell, G. Shaw, R. Swanstrom, M. Thomson, S. Tovanabutra, C. Williamson, and L. Zhang for contributing the HIV-1 Env plasmids used in our neutralization panel and N. Jean-Baptiste, R. Carroll, B. Flach, K. McKee, C. Moore, G. Padilla, J. Rathmann, S. O'Connell, S. O'Dell, S.D. Schmidt, C. Whittaker, and A.B. McDermott for assistance with neutralization assessments on the 208-strain panel. We thank J.D. Jou, A. Ojewole, and other members of the Donald lab for helpful discussions of antibody design strategies, and we thank A. Bartesaghi and J. Richardson for advice and guidance on the interpretation of cryo-EM densities. We thank Catherine Ehrhart for her invaluable help in producing models of antibody:Env complexes using AlphaFold. We gratefully acknowledge grant support from the NIH (R01 GM078031, R01 GM118543, and R35 GM144042 to B.R.D.). VRC research was funded by the Intramural Research Program of the Vaccine Research Center, NIAID, NIH. Some of this work was performed at the Simons Electron Microscopy Center and National Resource for Automated Molecular Microscopy located at the New York Structural Biology Center, supported by grants from the Simons Foundation (SF349247), NYSTAR, and the NIH National Institute of General Medical Sciences (GM103310).

AUTHOR CONTRIBUTIONS

N.A.D.-R., P.D.K., and B.R.D. designed research; J.G., S.W., A.U.L., B.Z., T.L., B.C.L., M.K.L., K.M., and S.O. performed research; G.T.H., J.G., M.S.F., R.R., and C.-H.S. analyzed data; and G.T.H., J.G., B.Z., P.D.K., and B.R.D. wrote the paper.

DECLARATION OF INTERESTS

B.R.D. and M.S.F. are founders of Ten63 Therapeutics. B.R.D., S.W., A.U.L., G.T.H., M.S.F., P.D.K., J.G., and N.A.D. are inventors on a patent application filed by Duke University.

Received: January 2, 2023
Revised: May 5, 2023
Accepted: June 12, 2023
Published: July 11, 2023

REFERENCES

1. Rudicell, R.S., Kwon, Y.D., Ko, S.Y., Pegu, A., Louder, M.K., Georgiev, I.S., Wu, X., Zhu, J., Boyington, J.C., Chen, X., et al. (2014). Enhanced Potency of a Broadly Neutralizing HIV-1 Antibody In Vitro Improves Protection against Lentiviral Infection In Vivo. J. Virol. *88*, 12669–12682. https://doi.org/10.1128/JVI.02213-14.
2. Julg, B., Tartaglia, L.J., Keele, B.F., Wagh, K., Pegu, A., Sok, D., Abbink, P., Schmidt, S.D., Wang, K., Chen, X., et al. (2017). Broadly neutralizing antibodies targeting the HIV-1 envelope V2 apex confer protection against a clade C SHIV challenge. Sci. Transl. Med. *9*, eaal1321. https://doi.org/10.1126/scitranslmed.aal1321.
3. Pegu, A., Borate, B., Huang, Y., Pauthner, M.G., Hessell, A.J., Julg, B., Doria-Rose, N.A., Schmidt, S.D., Carpp, L.N., Cully, M.D., et al. (2019). A Meta-analysis of Passive Immunization Studies Shows that Serum-Neutralizing Antibody Titer Associates with Protection against SHIV Challenge. Cell Host Microbe *26*, 336–346.e3. https://doi.org/10.1016/j.chom.2019.08.014.
4. Mahomed, S., Garrett, N., Baxter, C., Abdool Karim, Q., and Abdool Karim, S.S. (2021). Clinical Trials of Broadly Neutralizing Monoclonal Antibodies for Human Immunodeficiency Virus Prevention: A Review. J. Infect. Dis. *223*, 370–380. https://doi.org/10.1093/infdis/jiaa377.
5. Trkola, A., Kuster, H., Rusert, P., Joos, B., Fischer, M., Leemann, C., Manrique, A., Huber, M., Rehr, M., Oxenius, A., et al. (2005). Delay of HIV-1 rebound after cessation of antiretroviral therapy through passive transfer of human neutralizing antibodies. Nat. Med. *11*, 615–622. https://doi.org/10.1038/nm1244.
6. Ng, C.T., Jaworski, J.P., Jayaraman, P., Sutton, W.F., Delio, P., Kuller, L., Anderson, D., Landucci, G., Richardson, B.A., Burton, D.R., et al. (2010). Passive neutralizing antibody controls SHIV viremia and enhances B cell responses in infant macaques. Nat. Med. *16*, 1117–1119. https://doi.org/10.1038/nm.2233.
7. Caskey, M., Klein, F., Lorenzi, J.C.C., Seaman, M.S., West, A.P., Buckley, N., Kremer, G., Nogueira, L., Braunschweig, M., Scheid, J.F., et al. (2015). Viraemia suppressed in HIV-1-infected humans by broadly neutralizing antibody 3BNC117. Nature *522*, 487–491. https://doi.org/10.1038/nature14411.
8. Corey, L., Gilbert, P.B., Juraska, M., Montefiori, D.C., Morris, L., Karuna, S.T., Edupuganti, S., Mgodi, N.M., deCamp, A.C., Rudnicki, E., et al. (2021). Two Randomized Trials of Neutralizing Antibodies to Prevent HIV-1 Acquisition. N. Engl. J. Med. *384*, 1003–1014. https://doi.org/10.1056/NEJMoa2031738.
9. Georgiev, I.S., Doria-Rose, N.A., Zhou, T., Kwon, Y.D., Staupe, R.P., Moquin, S., Chuang, G.Y., Louder, M.K., Schmidt, S.D., Altae-Tran, H.R., et al. (2013). Delineating antibody recognition in polyclonal sera from patterns of HIV-1 isolate neutralization. Science *340*, 751–756. https://doi.org/10.1126/science.1233989.
10. Landais, E., Huang, X., Havenar-Daughton, C., Murrell, B., Price, M.A., Wickramasinghe, L., Ramos, A., Bian, C.B., Simek, M., Allen, S., et al. (2016). Broadly Neutralizing Antibody Responses in a Large Longitudinal Sub-Saharan HIV Primary Infection Cohort. PLoS Pathog. *12*, e1005369. https://doi.org/10.1371/journal. ppat.1005369.
11. Walker, L.M., Simek, M.D., Priddy, F., Gach, J.S., Wagner, D., Zwick, M.B., Phogat, S.K., Poignard, P., and Burton, D.R. (2010). A Limited Number of Antibody Specificities Mediate Broad and Potent Serum Neutralization in Selected HIV-1 Infected Individuals. PLoS Pathog. *6*, e1001028. https://doi.org/10.1371/journal.ppat.1001028.
12. Foley, B., Leitner, T., Apetrei, C., Hahn, B., Mirachi, I., Mullins, J., Rambaut, A., Wolinsky, S., and Korber, B. (2018). HIV Sequence Compendium 2018 Contracting Officer's Representative Los Alamos HIV Sequence Database and Analysis Staff. Technical Report. https://www.hiv.lanl.gov/.
13. Lee, J.H., Ozorowski, G., and Ward, A.B. (2016). Cryo-EM structure of a native, fully glycosylated, cleaved HIV-1 envelope trimer. Science *351*, 1043–1048. https://doi.org/10.1126/science.aad2450.
14. Stewart-Jones, G.B.E., Soto, C., Lemmin, T., Chuang, G.Y., Druz, A., Kong, R., Thomas, P.V., Wagh, K., Zhou, T., Behrens, A.J., et al. (2016). Trimeric HIV-1-Env Structures Define Glycan Shields from Clades A, B, and G. Cell *165*, 813–826. https://doi.org/10.1016/j.cell.2016.04.010.
15. Kwong, P.D., and Mascola, J.R. (2012). Human Antibodies that Neutralize HIV- 1: Identification, Structures, and B Cell Ontogenies. Immunity *37*, 412–425. https://doi.org/10.1016/j. immuni.2012.08.012.
16. Kwong, P.D., and Mascola, J.R. (2018). HIV-1 Vaccines Based on Antibody Identification, B Cell Ontogeny, and Epitope Structure. Immunity *48*, 855–871. https://doi.org/10.1016/j.immuni 2012.04.029.
17. McLellan, J.S., Pancera, M., Carrico, C., Gorman, J., Julien, J.P., Khayat, R., Louder, R., Pejchal, R., Sastry, M., Dai, K., et al. (2011). Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. Nature *480*, 336–343. https://doi.org/10.1038/nature10696.
18. Walker, L.M., Phogat, S.K., Chan-Hui, P.Y., Wagner, D., Phung, P., Goss, J.L., Wrin, T., Simek, M.D., Fling, S., Mitcham, J.L., et al. (2009). Broad and potent neutralizing antibodies from an african donor reveal a new HIV-1 vaccine target. Science *326*, 285–289. https://doi.org/10.1126/science.1178746.
19. Bonsignori, M., Hwang, K.K., Chen, X., Tsao, C.Y., Morris, L., Gray, E., Marshall, D.J., Crump, J.A., Kapiga, S.H., Sam, N.E., et al. (2011). Analysis of a Clonal Lineage of HIV-1 Envelope V2/V3 Conformational Epitope-Specific Broadly Neutralizing Antibodies and Their Inferred Unmutated

Common Ancestors. J. Virol. *85*, 9998–10009. https://doi.org/10.1128/jvi.05045-11.

20. Andrabi, R., Voss, J.E., Liang, C.H., Briney, B., McCoy, L.E., Wu, C.Y., Wong, C.H., Poignard, P., and Burton, D.R. (2015). Identification of Common Features in Prototype Broadly Neutralizing Antibodies to HIV Envelope V2 Apex to Facilitate Vaccine Design. Immunity *43*, 959–973. https://doi.org/10.1016/j.immuni.2015.10.014.
21. Gorman, J., Soto, C., Yang, M.M., Davenport, T.M., Guttman, M., Bailer, R.T., Chambers, M., Chuang, G.Y., DeKosky, B.J., Doria-Rose, N.A., et al. (2016). Structures of HIV-1 Env V1V2 with broadly neutralizing antibodies reveal commonalities that enable vaccine design. Nat. Struct. Mol. Biol. *23*, 81–90. https://doi.org/10.1038/nsmb.3144.
22. Walker, L.M., Huber, M., Doores, K.J., Falkowska, E., Pejchal, R., Julien, J.P., Wang, S.K., Ramos, A., Chan-Hui, P.Y., Moyle, M., et al. (2011). Broad neutralization coverage of HIV by multiple highly potent antibodies. Nature *477*, 466–470. https://doi.org/10.1038/nature10373.
23. Sok, D., van Gils, M.J., Pauthner, M., Julien, J.P., Saye-Francisco, K.L., Hsueh, J., Briney, B., Lee, J.H., Le, K.M., Lee, P.S., et al. (2014). Recombinant HIV envelope trimer selects for quaternary-dependent antibodies targeting the trimer apex. Proc. Natl. Acad. Sci. USA *111*, 17624–17629. https://doi.org/10.1073/pnas.1415789111.
24. Lee, J.H., Andrabi, R., Su, C.Y., Yasmeen, A., Julien, J.P., Kong, L., Wu, N.C., McBride, R., Sok, D., Pauthner, M., et al. (2017). A Broadly Neutralizing Antibody Targets the Dynamic HIV Envelope Trimer Apex via a Long, Rigidified, and Anionic β-Hairpin Structure. Immunity *46*, 690–702. https://doi.org/10.1016/j.immuni.2017.03.017.
25. Liu, Q., Acharya, P., Dolan, M.A., Zhang, P., Guzzo, C., Lu, J., Kwon, A., Gururani, D., Miao, H., Bylund, T., et al. (2017). Quaternary contact in the initial interaction of CD4 with the HIV-1 envelope trimer. Nat. Struct. Mol. Biol. *24*, 370–378. https://doi.org/10.1038/nsmb.3382.
26. Rantalainen, K., Berndsen, Z.T., Murrell, S., Cao, L., Omorodion, O., Torres, J.L., Wu, M., Umotoy, J., Copps, J., Poignard, P., et al. (2018). Co-evolution of HIV Envelope and Apex-Targeting Neutralizing Antibody Lineage Provides Benchmarks for Vaccine Design. Cell Rep. *23*, 3249–3261. https://doi.org/10.1016/j.celrep. 2018.05.046.
27. Gorman, J., Chuang, G.Y., Lai, Y.T., Shen, C.H., Boyington, J.C., Druz, A., Geng, H., Louder, M.K., McKee, K., Rawi, R., et al. (2020). Structure of Super-Potent Antibody CAP256-VRC26.25 in Complex with HIV-1 Envelope Reveals a Combined Mode of Trimer-Apex Recognition. Cell Rep. *31*, 107488. https://doi.org/10.1016/j.celrep.2020.03.052.
28. Pancera, M., Shahzad-Ul-Hussan, S., Doria-Rose, N.A., McLellan, J.S., Bailer, R.T., Dai, K., Loesgen, S., Louder, M.K., Staupe, R.P., Yang, Y., et al. (2013). Structural basis for diverse N-glycan recognition by HIV-1-neutralizing V1-V2-directed antibody PG16. Nat. Struct. Mol. Biol. *20*, 804–813. https://doi.org/10.1038/nsmb.2600.
29. Cale, E.M., Gorman, J., Radakovich, N.A., Crooks, E.T., Osawa, K., Tong, T., Li, J., Nagarajan, R., Ozorowski, G., Ambrozak, D.R., et al. (2017). Virus-like Particles Identify an HIV V1V2 Apex-Binding Neutralizing Antibody that Lacks a Protruding Loop. Immunity *46*, 777–791.e10. https://doi.org/10.1016/j.immuni.2017.04.011.
30. Barbas, C.F., Hu, D., Dunlop, N., Sawyer, L., Cababa, D., Hendry, R.M., Nara, P.L., and Burton, D.R. (1994). In vitro evolution of a neutralizing human antibody to human immunodeficiency virus type 1 to enhance affinity and broaden strain cross- reactivity. Proc. Natl. Acad. Sci. USA *91*, 3809–3813. https://doi.org/10.1073/pnas.91.9.3809.
31. Diskin, R., Scheid, J.F., Marcovecchio, P.M., West, A.P., Jr., Klein, F., Gao, H., Gnanapragasam, P.N.P., Abadir, A., Seaman, M.S., Nussenzweig, M.C., and Bjorkman, P.J. (2011). Increasing the potency and breadth of an HIV antibody by using structure-based rational design. Science *334*, 1289–1293. https://doi.org/10.1126/science.1213782.
32. Willis, J.R., Sapparapu, G., Murrell, S., Julien, J.P., Singh, V., King, H.G., Xia, Y., Pickens, J.A., LaBranche, C.C., Slaughter, J.C., et al. (2015). Redesigned HIV antibodies exhibit enhanced neutralizing potency and breadth. J. Clin. Invest. *125*, 2523–2531. https://doi.org/10.1172/JCI80693.
33. Wu, N.C., Grande, G., Turner, H.L., Ward, A.B., Xie, J., Lerner, R.A., and Wilson, I.A. (2017). In vitro evolution of an influenza broadly neutralizing antibody is modulated by hemagglutinin receptor specificity. Nat. Commun. *8*, 15371. https://doi.org/10.1038/ncomms15371.
34. Starr, T.N., Czudnochowski, N., Liu, Z., Zatta, F., Park, Y.J., Addetia, A., Pinto, D., Beltramello, M., Hernandez, P., Greaney, A.J., et al. (2021). SARS-CoV-2 RBD antibodies that maximize breadth and resistance to escape. Nature *597*, 7874. https://doi.org/10.1038/s41586-021-03807-6.
35. Rappazzo, C.G., Tse, L.V., Kaku, C.I., Wrapp, D., Sakharkar, M., Huang, D., Deveau, L.M., Yockachonis, T.J., Herbert, A.S., Battles, M.B., et al. (2021). Broad and potent activity against SARS-like viruses by an engineered human monoclonal antibody. Science *371*, 823–829. https://doi.org/10.1126/science.abf4830.
36. Lippow, S.M., Wittrup, K.D., and Tidor, B. (2007). Computational design of antibody-affinity improvement beyond in vivo maturation. Nat. Biotechnol. *25*, 1171–1176. https://doi.org/10.1038/nbt1336.
37. Sevy, A.M., Wu, N.C., Gilchuk, I.M., Parrish, E.H., Burger, S., Yousif, D., Nagel, M.B., Schey, K.L., Wilson, I.A., Crowe, J.E., Jr., et al. (2019). Multistate design of influenza antibodies improves affinity and breadth against seasonal viruses. Proc. Natl. Acad. Sci. USA *116*, 1597–1602. https://doi.org/10.1073/pnas.1806004116.
38. Tharakaraman, K., Robinson, L.N., Hatas, A., Chen, Y.L., Siyue, L., Raguram, S., Sasisekharan, V., Wogan, G.N., and Sasisekharan, R. (2013). Redesign of a cross-reactive antibody to dengue virus with broad-spectrum activity and increased in vivo potency. Proc. Natl. Acad. Sci. USA *110*, E1555–E1564. https://doi.org/10.1073/pnas.1303645110.
39. Barderas, R., Desmet, J., Timmerman, P., Meloen, R., and Casal, J.I. (2008). Affinity maturation of antibodies assisted by in silico modeling. Proc. Natl. Acad. Sci. USA *105*, 9029–9034. https://doi.org/10.1073/pnas. 0801221105.
40. Simonelli, L., Pedotti, M., Beltramello, M., Livoti, E., Calzolai, L., Sallusto, F., Lanzavecchia, A., and Varani, L. (2013). Rational Engineering of a Human Anti-Dengue Antibody through Experimentally Validated Computational Docking. PLoS One *8*, e55561. https://doi.org/10.1371/journal.pone.0055561.
41. Kratochvil, S., Shen, C.H., Lin, Y.C., Xu, K., Nair, U., Da Silva Pereira, L., Tripathi, P., Arnold, J., Chuang, G.Y., Melzi, E., et al. (2021). Vaccination in a humanized mouse model elicits highly protective PfCSP-targeting anti-malarial antibodies. Immunity *54*, 2859–2876.e7. https://doi.org/10.1016/j.immuni.2021.10.017.
42. Kwon, Y.D., Chuang, G.Y., Zhang, B., Bailer, R.T., Doria-Rose, N.A., Gindin, T.S., Lin, B., Louder, M.K., McKee, K., O'Dell, S., et al. (2018). Surface-Matrix Screening Identifies Semi-specific Interactions that Improve Potency of a Near Pan-reactive HIV-1-Neutralizing Antibody. Cell Rep. *22*, 1798–1809. https://doi.org/10.1016/j. celrep.2018.01.023.
43. Kwon, Y.D., Asokan, M., Gorman, J., Zhang, B., Liu, Q., Louder, M.K., Lin, B.C., McKee, K., Pegu, A., Verardi, R., et al. (2021). A matrix of structure-based designs yields improved VRC01- class antibodies for HIV-1 therapy and prevention. mAbs *13*. https://doi.org/10.1080/19420862.2021.1946918.
44. Hallen, M.A., Martin, J.W., Ojewole, A., Jou, J.D., Lowegard, A.U., Frenkel, M.S., Gainza, P., Nisonoff, H.M., Mukund, A., Wang, S., et al. (2018). OSPREY 3.0: Open-Source Protein Redesign for You, with Powerful New Features". J. Comput. Chem. *39*, 2494–2507. https://doi.org/10.1002/jcc.25522.
45. Lilien, R.H., Stevens, B.W., Anderson, A.C., and Donald, B.R. (2005). A novel ensemble-based scoring and search algorithm for protein redesign and its application to modify the substrate specificity of the gramicidin synthetase a phenylalanine adenylation enzyme. J. Comput. Biol. *12*, 740–761. https://doi.org/10.1089/cmb.2005.12.740.
46. Georgiev, I., Lilien, R.H., and Donald, B.R. (2008). The minimized dead-end elimination criterion and its application to p rotein redesign in a hybrid scoring and search algorithm for computing partition funct ions over molecular ensembles. J. Comput. Chem. *29*, 1527–1542. https://doi.org/10.1002/jcc.20909.

47. Donald, Bruce R. (2011). Algorithms in Structural Molecular Biology (The MIT Press).

48. Gainza, P., Roberts, K.E., and Donald, B.R. (2012). Protein design using continuous rotamers. PLoS Comput. Biol. *8*, e1002335. https://doi.org/10.1371/journal.pcbi.1002335.

49. Dennis, R.B., Williamson, R.A., and Parren, P.W.H.I. (2000). MINIREVIEW Antibody and Virus: Binding and Neutralization. Virology *270*, 1–3. https://doi.org/10.1006/viro.2000.0239.

50. Willis, J.R., Sapparapu, G., Murrell, S., Julien, J.P., Singh, V., King, H.G., Xia, Y., Pickens, J.A., LaBranche, C.C., Slaughter, J.C., et al. (2015). Redesigned HIV antibodies exhibit enhanced neutralizing potency and b readth. J. Clin. Invest. *125*, 2523–2531.

51. Ojewole, A.A., Jou, J.D., Fowler, V.G., and Donald, B.R. (2018). BBK* (Branch and Bound Over K*): A Provable and Efficient Ensemble-Based Protein Design Algorithm to Optimize Stability and Binding Affinity Over Large Sequence Spaces. J. Comput. Biol. *25*, 1557–8666. https://doi.org/10.1089/cmb.2017.0267.

52. Rawi, R., Mall, R., Shen, C.H., Farney, S.K., Shiakolas, A., Zhou, J., Bensmail, H., Chun, T.W., Doria-Rose, N.A., Lynch, R.M., et al. (2019). Accurate Prediction for Antibody Resistance of Clinical HIV- 1 Isolates. Sci. Rep. *9*, 14696. https://doi.org/10.1038/s41598-019-50635-w.

53. Breiman, L. (2001). Random Forests. Mach. Learn. *45*, 5–32.

54. Minoux, H., and Chipot, C. (1999). Cation-π interactions in proteins: Can simple models provide an accurate description? J. Am. Chem. Soc. *121*, 10366–10372. https://doi.org/10.1021/ja990914p.

55. Adolf-Bryfogle, J., Kalyuzhniy, O., Kubitz, M., Weitzner, B.D., Hu, X., Adachi, Y., Schief, W.R., and Dunbrack, R.L., Jr. (2018). RosettaAntibodyDesign (RAbD): A general framework for computational antibody design. PLoS Comput. Biol. *14*, e1006112. https://doi.org/10.1371/journal.pcbi.1006112.

56. Pantazes, R.J., and Maranas, C.D. (2010). OptCDR: a general computational method for the design of antibody complementarity determining regions for targeted epitope binding. Protein Eng. Des. Sel. *23*, 849–858. https://doi.org/10.1093/protein/gzq061.

57. Tong, L., Pantazes, R.J., and Maranas, C.D. (2014). OptMAVEn – A New Framework for the de novo Design of Antibody Variable Region Models Targeting Specific Antigen Epitopes. PLoS One *9*, e105954. https://doi.org/10.1371/journal.pone.0105954.

58. Chowdhury, R., Allan, M.F., and Maranas, C.D. (2018). OptMAVEn-2.0: De novo Design of Variable Antibody Regions Against Targeted Antigen Epitopes. Antibodies *7*, 23. https://doi.org/10.3390/antib7030023.

59. Lapidoth, G.D., Baran, D., Pszolla, G.M., Norn, C., Alon, A., Tyka, M.D., and Fleishman, S.J. (2015). AbDesign: An algorithm for combinatorial backbone design guided by natural conformations and sequences. Proteins *83*, 1385–1406. https://doi.org/10.1002/prot.24779.

60. Ruffolo, J.A., Sulam, J., and Gray, J.J. (2022). Antibody structure prediction using interpretable deep learning. Patterns *3*, 100406. https://doi.org/10.1016/j.patter.2021.100406.

61. Alexander, M. (2015). Sevy et al. ''Design of Protein Multi-specificity Using an Independent Sequence Search Reduces the Barrier to Low Energy Sequences''. PLoS Comput. Biol. *11*, e1004300. https://doi.org/10.1371/journal.pcbi.1004300.

62. Rawi, R., Rutten, L., Lai, Y.T., Olia, A.S., Blokland, S., Juraszek, J., Shen, C.H., Tsybovsky, Y., Verardi, R., Yang, Y., et al. (2020). Automated Design by Structure-Based Stabilization and Consensus Repair to Achieve Prefusion-Closed Envelope Trimers in a Wide Variety of HIV Strains. Cell Rep. *33*, 108432. https://doi.org/10.1016/j.celrep. 2020.108432.

63. Chuang, G.Y., Zhou, J., Acharya, P., Rawi, R., Shen, C.H., Sheng, Z., Zhang, B., Zhou, T., Bailer, R.T., Dandey, V.P., et al. (2019). Structural Survey of Broadly Neutralizing Antibodies Targeting the HIV-1 Env Trimer Delineates Epitope Categories and Characteristics of Recognition. Structure *27*, 196–206.e6. https://doi.org/10.1016/j.str.2018.10.007.

64. Chen, V.B., Davis, I.W., and Richardson, D.C. (2009). KiNG (Kinemage, Next Generation): A versatile interactive molecular and scientific visualization program. Protein Sci. *18*, 2403–2409. https://doi.org/10.1002/pro.250.

65. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera - A visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612. https://doi.org/10.1002/jcc.20084.

66. Suloway, C., Pulokas, J., Fellmann, D., Cheng, A., Guerra, F., Quispe, J., Stagg, S., Potter, C.S., and Carragher, B. (2005). Automated molecular microscopy: The new Leginon system. J. Struct. Biol. *151*, 41–60. https://doi.org/10.1016/j.jsb.2005.03.010.

67. Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods *14*, 290–296. https://doi.org/10.1038/nmeth.4169.

68. Barad, B.A., Echols, N., Wang, R.Y.R., Cheng, Y., DiMaio, F., Adams, P.D., and Fraser, J.S. (2015). EMRinger: side chain–directed model and map validation for 3D cryo-electron microscopy. Nat. Methods *12*, 943–946. https://doi.org/10.1038/nmeth.3541.

69. Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Cryst *66*, 213–221. https://doi.org/10.1107/s0907444909052925.

70. Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Cryst *60*, 2126–2132. https://doi.org/10.1107/s0907444904019158.

71. Williams, C.J., Headd, J.J., Moriarty, N.W., Prisant, M.G., Videau, L.L., Deis, L.N., Verma, V., Keedy, D.A., Hintze, B.J., Chen, V.B., et al. (2018). MolProbity: More and better reference data for improved all-atom structure validation. Protein Sci. *27*, 293–315. https://doi.org/10.1002/pro.3330.

72. Lowegard, A.U., Frenkel, M.S., Holt, G.T., Jou, J.D., Ojewole, A.A., and Donald, B.R. (2020). Novel, provable algorithms for efficient ensemble-based computational protein design and their application to the redesign of the c-Raf- RBD:KRas protein-protein interface. PLoS Comput. Biol. *16*, e1007447. https://doi.org/10.1371/journal.pcbi.1007447.

73. Julien, J.P., Lee, J.H., Cupo, A., Murin, C.D., Derking, R., Hoffenberg, S., Caulfield, M.J., King, C.R., Marozsan, A.J., Klasse, P.J., et al. (2013). Asymmetric recognition of the HIV-1 trimer by broadly neutralizing antibody PG9. Proc. Natl. Acad. Sci. USA *110*, 4351–4356. https://doi.org/10.1073/pnas.1217537110.

74. Krieger, E., Joo, K., Lee, J., Lee, J., Raman, S., Thompson, J., Tyka, M., Baker, D., and Karplus, K. (2009). Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: four approaches that performed well in CASP8. Proteins: Struct., Funct., Bioinf. *77*, 114–122.

75. Word, J.M., Lovell, S.C., Richardson, J.S., and Richardson, D.C. (1999). Asparagine and glutamine: Using hydrogen atom contacts in the choice of side-chain amide orientation. J. Mol. Biol. *285*, 1735–1747. https://doi.org/10.1006/jmbi.1998.2401.

76. Lovell, S.C., Word, J.M., Richardson, J.S., and Richardson, D.C. (2000). The penultimate rotamer library. Proteins *40*, 389–408.

77. Wang, S. (2021). Computational Protein Design with Non-proteinogenic Amino Acids and Small Molecule Ligands, with Applications to Protein-Protein Interaction Inhibitors, Anti-microbial Enzyme Inhibitors, and Antibody Design (Duke University). PhD thesis.

78. Case, D.A., Darden, T.A., Cheatham, T.E., III, Simmerling, C.L., Wang, J., Duke, R.E., Luo, R., Merz, K.M., Pearlman, D.A., Crowley, M., et al. (2006). AMBER 9 (University of California), p. 45.

79. Sarzotti-Kelsoe, M., Bailer, R.T., Turk, E., Lin, C.L., Bilska, M., Greene, K.M., Gao, H., Todd, C.A., Ozaki, D.A., Seaman, M.S., et al. (2014). Optimization and validation of the TZM-bl assay for standardized assessments

of neutralizing antibodies against HIV-1. J. Immunol. Methods *409*, 131–146. https://doi.org/10.1016/j.jim.2013.11.022.

80. Wu, X., Yang, Z.Y., Li, Y., Hogerkorp, C.M., Schief, W.R., Seaman, M.S., Zhou, T., Schmidt, S.D., Wu, L., Xu, L., et al. (2010). Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. Science *329*, 856–861. https://doi.org/10.1126/science.1187659.
81. Pedregosa, F., Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., et al. (2011). Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res. *12*, 2825–2830.
82. Cock, P.J., Antao, T., Chang, J.T., Chapman, B.A., Cox, C.J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., et al. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics *25*, 1422–1423. https://doi.org/10.1093/bioinformatics/btp163.
83. Friedman, J.H. (2001). Greedy function approximation: A gradient boosting machine. Ann. Statist. *29*, 1189–1232. https://doi.org/10.1214/aos/1013203451.
84. Hunter, J.D. (2007). Matplotlib: A 2D graphics environment. Comput. Sci. Eng. *9*, 90–95. https://doi.org/10.1109/MCSE.2007.55.
85. Waskom, Michael L. (2021). “seaborn: statistical data visualization”. J. Open Source Softw. *6*, 3021. https://doi.org/10.21105/joss.03021.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
PG9RSH-(DU001 – DU010, N(100fY), DU012 – DU034)	NIH/VRC, this paper	N/A
PGT145-(DU301 – DU310)	NIH/VRC, this paper	N/A
Monoclonal anti-HIV-1 Env PGT145	NIH AIDS Reagent Program[22]	CAT# 12703, RRID: AB_2491054
Monoclonal anti-HIV-1 Env PG9-16-RSH	NIH/VRC[28]	N/A
Monoclonal anti-HIV-1 Env PGDM1400	Dennis R. Burton, Scripps[23]	N/A
Bacterial and virus strains		
VRC 208 virus panel	NIH/VRC[63]	N/A
VRC 10 virus panel (PGT145)	NIH/VRC	N/A
VRC 10 virus panel (PG9RSH)	NIH/VRC	N/A
Chemicals, peptides, and recombinant proteins		
BG505 DS-SOSIP.664	NIH/VRC	N/A
Critical commercial assays		
Turbo293 Ab Transfection Kit	Speed Biosystems	Cat# PXX1005
Human Antibody Capture Kit	GE Healthcare Life Sciences	Cat# BR-1008-39
Deposited data		
Cryo-EM map: PG9RSH N(100f)Y in complex with BG505 Env Trimer	EMDB	29248
Cryo-EM structure: PG9RSH N(100f)Y in complex with BG505 Env Trimer	PDB	8FK5
Cryo-EM map: PG9RSH DU025 in complex with BG505 Env Trimer	EMDB	29264
Cryo-EM structure: PG9RSH DU025 in complex with BG505 Env Trimer	PDB	8FL1
Cryo-EM map: PGT145 DU303 in complex with BG505 Env Trimer	EMDB	29288
Cryo-EM structure: PGT145 DU303 in complex with BG505 Env Trimer	PDB	8FLW
Code required for replication of designs and analysis	Harvard Dataverse	https://doi.org/10.7910/DVN/NXD2JR
Experimental models: Cell lines		
TZM-bl cells	NIH AIDS Reagent Program	Cat# 8129
Expi293F cells	ThermoFisher Scientific Inc	Cat# A14527
Recombinant DNA		
pVRC8400	https://www.addgene.org/	Cat# 63160
pVRC8400-PG9RSH-(DU001 – DU034) Heavy	NIH/VRC, this paper	N/A
pVRC8400-PG9RSH-(DU001 – DU034) Light	NIH/VRC, this paper	N/A
pVRC8400-PGT145-(DU301 – DU310) Heavy	NIH/VRC, this paper	N/A
pVRC8400-PGT145-(DU301 – DU310) Light	NIH/VRC, this paper	N/A
Software and algorithms		
OSPREY	Hallen et al.[44]	https://github.com/donaldlab/OSPREY3
The PyMol Molecular Graphics System	Schrödinger, LLC	https://pymol.org/2/
KING	Chen et al.[64]	https://github.com/rlabduke/javadev
USCF Chimera	Pettersen et al.[65]	https://www.cgl.ucsf.edu/chimera/
Leginon	Suloway et al.[66]	https://sbgrid.org/software/titles/leginon

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
cryoSPARC	Punjani et al.[67]	https://cryosparc.com/
EMRinger	Barad et al.[68]	https://github.com/fraser-lab/EMRinger
Phenix	Adams et al.[69]	https://sbgrid.org/software/
Coot	Emsley and Cowtan.[70]	https://sbgrid.org/software/
MolProbity	Williams et al.[71]	http://molprobity.biochem.duke.edu/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Bruce R. Donald (brd+cell22@cs.duke.edu).

Materials availability

TZM-bl cells (https://www.aidsreagent.org, cat# 8129) are available through the NIH HIV Reagent Program, Division of AIDS, NIAID, NIH, contributed by Dr. John C. Kappes, Dr. Xiaoyun Wu and Tranzyme Inc.

Data and code availability

Neutralization data are included in this manuscript in the supplemental information. Cryo-EM maps and structures have been deposited in the EMDB (EMDB: EMD-299248, EMD-29264, EMD-29288) and the PDB (PDB: 8FK5, 8FL1, 8FLW), respectively, and are publicly available as of the date of publication.

All code used and discussed in this manuscript is available from the Harvard Dataverse repository (Harvard Dataverse: https://doi.org/10.7910/DVN/NXD2JR) as of the date of publication. For new empirical designs we recommend using the latest version of OSPREY available for free at http://www.cs.duke.edu/donaldlab/opsrey.php. All computer code for the OSPREY system is also available on GitHub at https://github.com/donaldlab/OSPREY3, and is open/source and free. The version of record is archived with Zenodo (Zenodo: https://doi.org/10.5281/zenodo.7633931).

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

These cells are a HeLa cell line generated from JC.53 cells that expresses CD4, CCR5, and CXCR4, with galactosidase and luciferase reporter genes under the HIV-1 promoter. For long-term storage store at or below -100° C, preferably in FBS supplemented with 40% DMEM and 10% DMSO. Propagate in DMEM supplemented with 10% FBS, 100 U per mL penicillin and 0.1 mg per mL streptomycin, incubate at 37° C. For more information see the NIH AIDS Reagent Program (https://www.aidsreagent.org, cat# 8129).

METHOD DETAILS

Redesign of PG9RSH and PGT145

Designs to obtain improved variants of PG9RSH and PGT145 were performed by defining sets of accessible conformations (conformation spaces) for unliganded antibody, unliganded Env trimer, and complexed antibody:Env states, followed by approximation of binding affinity using the K^* algorithm[45,46,48] or an early version of the $EWAK^*$ algorithm[72] in OSPREY.

Conformation spaces were defined for PG9RSH or PGT145 based on an EM structure of the PG9 and 8ANC195 bNAbs in complex with the BG505 SOSIP.664 Env trimer (PDB: 5VJ6)[73] or a Cryo-EM structure of PGT145 and sCD4 in complex with the DS-SOSIP.664 (Based on PDB: 5U1F),[25] respectively. The structure of PGT145 contained modeled locations of amino acid side chains and glycans, which were not present in the deposited structure due to resolution limitations.

Structures were inspected to determine whether they were suitable for design or needed refinement. MolProbity[71] analysis of the CDRH3 region of PG9 (5VJ6) revealed a few major clashes, one of which involved antibody backbone atoms, indicating that the atomic model may represent an inaccurate backbone conformation. As a result, the PG9 input structure was all-atom minimized using Yasara[74] to relax steric clashes, and the 8ANC195 antibody was removed along with some distal regions of gp160. This step was performed due to the very low resolution of the structure, coupled with the evident backbone clashes. Similar analysis of the PGT145 structure based on 5U1F (with added side chains) showed only a few clashes, each of which involved only side chain conformations. We chose not to perform all-atom minimization of this structure, because we did not consider the relaxation of side chain conformations to be worth changing the backbone conformation without experimental evidence. The PGT145 input structure was modified by removing sCD4 along with parts of the PGT145 antibody and gp160 distal from peptide contacts of the PGT145:Env interface.

Hydrogens were added to both input structures using Reduce.[75] Backbone coordinates for the complex were defined by the resulting modified PG9:Env and PGT145:Env structures, and coordinates for the unliganded antibody and unliganded Env states were obtained by removing atoms corresponding to the Env and antibody, respectively. Design residues (See Table S1) were modeled as continuously flexible[48] in OSPREY, for which rotamers from the Penultimate Rotamer Library[76] were allowed to adopt any side-chain conformation such that all χ-angles are within $\pm 9^{\circ}$ of their modal χ-angles. All other side-chain coordinates were obtained from the input structures. Rotamers and energetic interactions for sulfated tyrosines were modeled using the methods reported in Wang.[77] Briefly, a rotamer library was constructed, partial charges and force-field parameters were computed with Antechamber in AMBER,[78] and solvation parameters were computed using an extended version of the EEF1 solvation model. For each model we computed ε-approximate bounds on the K^* score to a guaranteed accuracy[51] of $\varepsilon < 0.683$ using the K^*[45,46,48] or $EWAK^*$[72] algorithms.

Antibody variant expression and purification

DNA sequences of heavy and light chain variable regions of antibodies PG9RSH and PGT145 and variants were synthesized and subcloned into the pVRC8400 vector. For antibody expression, equal amounts of antibody heavy and light chain plasmid DNA were transfected into Expi293 cells using Turbo293 transfection reagent (Speed BioSystems). The transfected cells were incubated in shaker incubator at 120 rpm, 37° C, 9% CO_2. The culture supernatants were harvested, filtered, and loaded on a protein A (GE Healthcare) column at 5 days post transfection. After washing the column with PBS, each antibody was eluted with an IgG elution buffer (Pierce) and immediately neutralized with one-tenth volume of 1M Tris-HCl pH 8.0. Eluted antibodies were dialyzed against PBS overnight and were confirmed by SDS-PAGE before use.

Pseudovirus neutralization assays

Antibody neutralization was evaluated with the single-round infection assay of TZM-bl cells.[79] Antibodies were serially diluted into wells of a 384-well plate, a constant amount of pseudovirus was added, plates were incubated for 60 min, and TZM-bl cells, which cells express luciferase upon viral infection, were added. Plates were incubated for 48 h, lysed, and measured for luciferase activity. The antibody concentration required to achieve 50% neutralization of infection (IC_{50}) was calculated using a dose-response curve fit with a 5-parameter nonlinear function. For small-panel neutralization assays we used a panel of 10 HIV-1 Env pseudoviruses from clades A, B, and C. For large-panel neutralization assays we used a previously described panel[9,29,80] of 208 geographically and genetically diverse HIV-1 Env pseudoviruses representing the major subtypes and circulating recombinant forms. All IC_{50} values reported here are from small (10 viruses) or large (complete set of 208 viruses) neutralization panels run at the VRC. In some cases, multiple runs were averaged. We report both the potency (measured as the median or geometric mean IC_{50} or IC_{80} for strains with measurable neutralization) and the breadth of neutralization (the number or percentage of strains with measurable neutralization). These summary statistics were computed in this way to conform to the literature standard and to enable straightforward comparison. Sources of error include the fact that neutralization IC_{50} values are known to vary up to 3-fold between repeat assays.[80]

Cryo-EM data collection, structure determination, and refinement

The BG505 DS-SOSIP.664 Env trimer was incubated with molar excess of antigen-binding fragment (Fab) for each of the improved V2-apex directed antibodies. Grids were prepared by depositing 2 μL of each complex at 2 mg/mL final concentration on C-flat 1.2/1.3 grids (emsdiasum.com) and vitrified with an FEI Vitrobot Mark IV with a wait time of 30 s, blot time of 3 s, and blot force of 1. Data collections were performed on a Titan Krios electron microscope with Leginon using a Gatan K3 direct detection device. Exposures were collected in movie mode for 2 s with the total dose of 63.75 $e^{\circ -}/\mathring{A}^{2}$ fractionated over 40 raw frames. cryoSPARC v3.1 was used for frame alignment, CTF estimation, 2D classifications, ab initio 3D reconstruction, homogeneous refinement, and nonuniform 3D refinement. 3D reconstruction and final refinements were performed using C1 symmetry.

Coordinates from PDB: 5V8L and PDB: 3U4E were used for initial fits to the reconstructed maps. This was followed by simulated annealing and real space refinement in Phenix v1.19 with the sharpened map from cryoSPARC v3.1 and with a density modified map from Phenix Resolve and manually fit with Coot v0.9.8 and then improved through iterative rounds. Geometry and map fitting parameters were evaluated using Molprobity v4.5.1 and EMRinger. Maps and structures were deposited to the EMDB (EMDB: EMD-29248, EMD-29264, EMD-29288) and PDB (PDB: 8FK5, 8FL1, 8FLW).

QUANTIFICATION AND STATISTICAL ANALYSIS

Predicting change in neutralization from Env sequence

We modified a previously reported method for predicting bNAb neutralization from Env sequence[52] to predict the consequences of bNAb single mutations (increased or decreased neutralization). We constructed models using gradient-boosted decision trees in scikit-learn,[81] which uses a boosting approach to construct ensemble models of CART decision trees.

To define labels corresponding to the change in neutralization relative to wild-type for each antibody we computed the log-ratio of neutralization for each antibody (DU303, PG9RSH N(100f)Y, DU025) and its corresponding ancestor (PGT145, PG9RSH): $z = \log_{10}\frac{IC_{50}^{WT}}{IC_{50}^{mut}}$. We defined labels y for binary classification where $y = 1$ if $z > 0$, $y = 0$ otherwise. We processed Env protein alignments using BioPython.[82] To generate Env sequence features *X* we first augmented the Env protein sequences by identifying potential N-glycosylation sites, defined as sites containing the amino acid motif N-X-S/T, where X represents any amino acid. This resulted in 957 categorical features with an alphabet size of 21. Final features were obtained by one-hot encoding, resulting in a total of 4939 binary features.

For training we optimized three hyperparameters, leaving the rest at default values. We used an early stopping criterion implemented in sci-kit learn for training: 10% of the training data was held out as an additional validation set, and training was halted if the score on the validation set did not improve for a user-specified number of iterations. We optimized the maximum depth of the CART decision trees in the ensemble, the ”learning rate” - a scaling of the contribution of each decision tree to the overall decision function, and the number of iterations of no improvement required for the early stopping criterion. Hyperparameters were optimized by 10-fold cross-validation (repeated 5 times) and parameters were selected by computing the average accuracy, AUC, or F1 score on the validation set.

Variable importance, measured by mean decrease in impurity (MDI) and permutation importance (PI), was evaluated for DU303 on a model trained using the entire available dataset. The MDI variable importance measure is analogous to the Gini importance - for each feature its MDI is defined as the average decrease in impurity over all nodes that correspond to the feature.[83] In this case our splitting criterion is the Friedman Mean-Squared Error, Equation 35 in Friedman.[83] The MDI importance was computed using the scikit-learn implementation. PI was computed by randomly permuting each feature and then computing the difference in loss between using scrambled and original features using the scikit-learn implementation (`sklearn.inspection.permutation_importance`).

Visualization and figure generation

Structure and density was visualized using PyMOL, USCF Chimera,[65] and King,[64] and images were generated with PyMOL. Analysis of neutralization data was performed using Python, and accompanying figures were generated using the Matplotlib[84] and Seaborn[85] libraries.

Cell Reports, Volume *42*

Supplemental information

Improved HIV-1 neutralization breadth and potency of V2-apex antibodies by *in silico* design

Graham T. Holt, Jason Gorman, Siyu Wang, Anna U. Lowegard, Baoshan Zhang, Tracy Liu, Bob C. Lin, Mark K. Louder, Marcel S. Frenkel, Krisha McKee, Sijy O'Dell, Reda Rawi, Chen-Hsiang Shen, Nicole A. Doria-Rose, Peter D. Kwong, and Bruce R. Donald

Table S1: Flexible residues for designs of PG9RSH and PGT145, related to STAR Methods.

PDB ID	**Name**	**Target**	**Mutable res**	**Flexible res**
5VJ6	PG9:Env	100l	H1839	E953, F1418, H1835, H1841
5VJ6	PG9:Env	100i, 100j, 100k	H1836, H1837, H1838	E955, E956, E957, E958, E1091, H1754, H1776, H1842
5VJ6	PG9:Env	100, 100b, 100c	H1827, H1829, H1830	D503, D505, E945, H1834
5VJ6	PG9:Env	100e	H1832	E951, E952, E954, F1403
5VJ6	PG9:Env	100f	H1833	E952, E953, H1828, H1835
5VJ6	PG9:Env	100k	H1838	E955, E1089, E1091, E1092, H1837
5U1F	PGT145:Env	100l	H230	A14, D146, H222, H224, H232
5U1F	PGT145:Env	100d, 100l	H222, H230	A15, D145, D146, H219, H232
5U1F	PGT145:Env	100e, 100m	H223, H231	C84, C87, D143, H225

Table S2: Upper and lower bounds on the K* score and partition functions for PGT145 designs at residues 100d and 100l predicted using OSPREY, related to Results. H_δ and H_ε refer to the δ and ε protonation states of histidine.

		$\log_{10}$K*		$\log_{10} Z_{\text{C}}$		$\log_{10} Z_{\text{Ab}}$		$\log_{10} Z_{\text{Env}}$	
100d	100l	UB	LB	UB	LB	UB	LB	UB	LB
E	D	10.652	9.599	30.550[1]	30.055	12.802[1]	12.307	8.086[1]	7.591
E	E	10.020	8.564	29.780[1]	29.285	12.667[1]	12.172	8.082[1]	7.588
D	D	9.882	8.431	30.335[1]	29.841	13.384[1]	12.889	8.059[1]	7.564
D	E	9.077	8.108	29.555[1]	29.060	13.380[1]	12.885	8.087[1]	7.592
E	Y	8.980	8.150	26.480[1]	25.985	10.284[1]	9.789	8.206[1]	7.711
F	D	8.891	7.440	35.516[1]	35.021	19.540[1]	19.045	8.074[1]	7.579
D	Y	8.704	7.744	26.486[1]	25.991	10.679[1]	10.184	8.092[1]	7.597
H_δ	D	8.644	7.236	34.024	33.529	18.234	17.777	8.059	7.603
M	D	8.537	7.197	31.473	30.978	15.742	15.288	8.040	7.648
E	H_ϵ	8.350	7.491	31.993[1]	31.498	16.468[1]	15.973	8.164[1]	7.669
W	Y	7.933	6.679	29.092	28.598	13.862	13.555	8.058	7.604
H_ϵ	D	7.926	6.878	33.248[1]	32.753	18.206[1]	17.711	8.105[1]	7.610

Upper and lower bounds on the K* score and partition functions for PGT145 designs at residues 100d and 100l predicted using OSPREY (continued).

		$\log_{10}$K*		$\log_{10} Z_{\mathrm{C}}$		$\log_{10} Z_{\mathrm{Ab}}$		$\log_{10} Z_{\mathrm{Env}}$	
100d	100l	UB	LB	UB	LB	UB	LB	UB	LB
F	E	7.804	6.829	33.995[1]	33.500	19.099[1]	18.604	8.081[1]	7.586
D	H_ϵ	7.665	6.815	31.919[1]	31.424	17.072[1]	16.577	8.171[1]	7.676
Y	D	7.622	6.430	32.578	32.083	17.597	17.392	8.056	7.564
H_δ	E	7.532	6.224	32.785	32.290	18.049	17.555	8.017	7.699
W	D	7.482	6.490	33.303	32.808	18.263	18.254	8.055	7.568
M	E	7.459	6.096	30.804	30.309	16.161	15.788	8.051	7.557
E	Q	7.144	6.053	30.916[1]	30.421	16.692[1]	16.197	8.070[1]	7.575
E	M	7.025	6.015	31.697[1]	31.202	17.539[1]	17.044	8.123[1]	7.628
H_ϵ	E	6.969	5.810	31.567[1]	31.072	17.504[1]	17.009	8.084[1]	7.589
Y	H_ϵ	6.961	5.543	31.758	31.264	17.674	17.185	8.046	7.612
F	H_ϵ	6.933	5.965	34.419[1]	33.924	20.380[1]	19.885	8.096[1]	7.601
E	N	6.911	5.609	32.204[1]	31.709	18.159[1]	17.665	8.123[1]	7.629
H_δ	Y	6.907	5.477	28.426	27.931	14.394	13.953	8.060	7.566
W	E	6.833	5.796	32.123	31.628	17.790	17.642	8.042	7.648
W	H_ϵ	6.809	5.461	34.234	33.740	20.221	19.856	8.058	7.568
M	Y	6.786	5.834	26.016	25.521	11.647	11.647	8.040	7.583
Y	Y	6.747	5.758	27.386	26.891	13.079	13.079	8.054	7.561
E	V	6.719	5.804	32.329[1]	31.834	18.473[1]	17.978	8.126[1]	7.631
H_δ	H_ϵ	6.700	5.485	33.744	33.250	19.733	19.394	8.032	7.650
E	R	6.689	5.441	34.344[1]	33.849	20.557[1]	20.062	8.087[1]	7.593
F	Y	6.680	5.804	29.784[1]	29.289	15.940[1]	15.445	8.154[1]	7.659
E	S	6.666	5.206	31.168[1]	30.673	17.401[1]	16.906	8.091[1]	7.596
E	C	6.633	5.222	31.129[1]	30.634	17.391[1]	16.897	8.094[1]	7.599
E	T	6.605	5.197	32.034[1]	31.540	18.292[1]	17.798	8.126[1]	7.631
M	H_ϵ	6.454	5.332	31.439	30.944	17.591	17.292	8.021	7.692
D	Q	6.404	5.050	30.719[1]	30.224	17.228[1]	16.733	8.077[1]	7.582
D	M	6.381	5.071	31.511[1]	31.016	18.049[1]	17.554	8.071[1]	7.576
D	N	6.282	4.802	31.922[1]	31.427	18.561[1]	18.066	8.069[1]	7.574
H_ϵ	Y	6.257	5.285	27.469	26.974	13.636	13.636	8.053	7.575
D	R	6.114	4.739	33.798[1]	33.303	20.581[1]	20.086	8.093[1]	7.598
H_ϵ	H_ϵ	6.049	4.582	32.698	32.203	19.562	19.083	8.060	7.566
D	V	6.004	4.883	32.076[1]	31.582	18.978[1]	18.483	8.084[1]	7.589
D	C	5.976	4.529	30.901[1]	30.406	17.843[1]	17.348	8.072[1]	7.577
D	S	5.945	4.471	31.010[1]	30.515	17.978[1]	17.484	8.076[1]	7.581
D	T	5.813	4.518	31.841[1]	31.346	18.833[1]	18.338	8.184[1]	7.690
Y	N	5.467	4.200	32.817	32.322	20.060	19.772	8.061	7.577
H_δ	M	5.377	4.343	33.583	33.089	20.724	20.533	8.021	7.673
F[2]	N[2]	5.288	3.883	35.738[1]	35.243	23.342[1]	22.847	8.098[1]	7.603
H_δ	R	5.238	3.981	34.091	33.597	21.558	21.280	8.058	7.574
F	V	5.213	3.917	35.501[1]	35.006	23.206[1]	22.711	8.072[1]	7.577
M	R	5.120	3.825	32.709	32.214	20.329	20.023	8.060	7.566
Y	S	5.054	3.579	34.155	33.660	22.016	21.524	8.066	7.577
M	M	5.044	3.602	31.930	31.436	19.785	19.313	8.048	7.574

Upper and lower bounds on the K* score and partition functions for PGT145 designs at residues 100d and 100l predicted using OSPREY (continued).

		$\log_{10}$K*		$\log_{10} Z_{\text{C}}$		$\log_{10} Z_{\text{Ab}}$		$\log_{10} Z_{\text{Env}}$	
100d	100l	UB	LB	UB	LB	UB	LB	UB	LB
M	Q	5.019	3.751	30.786	30.292	18.502	18.149	8.039	7.618
M	N	5.007	3.673	31.933	31.438	19.734	19.241	8.031	7.685
H$_\delta$	N	4.990	3.670	34.513	34.018	22.327	21.843	8.022	7.680
H$_\delta$	V	4.928	3.474	34.630	34.135	22.611	22.119	8.050	7.583
Y	C	4.923	3.659	31.450	30.955	19.244	18.927	8.052	7.600
F	Q	4.921	3.491	34.407[1]	33.912	22.401[1]	21.906	8.076[1]	7.581
F	T	4.914	3.507	35.738[1]	35.243	23.698[1]	23.204	8.115[1]	7.621
F	S	4.914	3.434	34.748[1]	34.253	22.752[1]	22.257	8.072[1]	7.577
F	C	4.871	3.570	34.570[1]	34.075	22.489[1]	21.994	8.199[1]	7.704
M	V	4.864	3.885	32.351	31.856	19.918	19.918	8.053	7.569
H$_\delta$	C	4.857	3.384	33.457	32.963	21.518	21.032	8.061	7.568
H$_\delta$	Q	4.817	3.462	33.203	32.708	21.191	20.819	8.056	7.567
M	T	4.770	3.296	31.964	31.469	20.115	19.626	8.057	7.568
M	C	4.754	3.615	30.973	30.479	18.804	18.653	8.060	7.566
H$_\delta$	T	4.748	3.280	34.548	34.053	22.717	22.228	8.057	7.572
H$_\delta$	S	4.736	3.290	33.578	33.083	21.741	21.246	8.052	7.596
W	N	4.733	3.398	33.926	33.431	21.975	21.579	8.058	7.614
H$_\epsilon$	M	4.714	3.309	32.655	32.161	20.809	20.378	8.043	7.564
M	S	4.633	3.313	31.073	30.579	19.241	18.746	8.024	7.694
W	Q	4.575	3.609	32.884	32.389	20.730	20.722	8.050	7.587
W	V	4.544	3.542	34.654	34.160	22.564	22.514	8.054	7.596
H$_\epsilon$	N	4.511	3.031	33.192[1]	32.698	21.601[1]	21.106	8.071[1]	7.576
W	T	4.363	2.902	34.169	33.674	22.717	22.234	8.056	7.571
W	S	4.326	2.848	33.083	32.588	21.687	21.196	8.053	7.561
W	C	4.299	2.995	32.965	32.471	21.451	20.960	8.025	7.706
H$_\epsilon$	Q	4.294	2.860	32.016	31.521	20.600	20.150	8.061	7.571
H$_\epsilon$	R	4.294	2.989	32.219[1]	31.724	20.829[1]	20.334	8.085[1]	7.590
H$_\epsilon$	V	4.293	2.860	33.453[1]	32.958	22.069[1]	21.574	8.081[1]	7.586
H$_\epsilon$	C	4.198	2.802	32.159	31.664	20.804	20.392	8.058	7.569
Y	E	4.144	3.041	27.191	26.696	15.594	15.454	8.062	7.593
H$_\epsilon$	S	4.123	2.641	32.422[1]	31.927	21.222[1]	20.727	8.067[1]	7.572
H$_\epsilon$	T	4.118	2.650	33.337	32.842	22.127	21.632	8.066	7.586
F	R	3.277	1.925	31.262[1]	30.768	20.882[1]	20.387	8.093[1]	7.598
F	M	3.164	2.066	30.575[1]	30.080	20.200[1]	19.705	8.201[1]	7.706
Y	Q	2.212	0.901	26.936	26.441	17.481	17.127	8.059	7.597
Y	T	2.175	0.691	28.388	27.893	19.144	18.650	8.058	7.564
Y	R	2.046	0.734	28.979	28.484	19.716	19.254	8.033	7.679
Y	M	2.024	0.685	28.333	27.839	19.113	18.644	8.040	7.665
Y	V	1.915	0.578	28.362	27.868	19.259	18.769	8.031	7.679
W	M	1.180	−0.220	30.092	29.597	21.768	21.283	8.048	7.629
W	R	−4.358	−5.607	24.388	23.894	21.470	21.050	8.031	7.696

[1] Upper bounds computed using lower bound and $\epsilon = 0.68$.

[2] Wild-type amino-acid.

Table S3: Upper and lower bounds on the K^* score and partition functions for PG9 designs at residue 100f predicted using OSPREY, related to Results. H_δ and H_ε refer to the δ and ε protonation states of histidine.

	$\log_{10} K^*$		$\log_{10} Z_{\mathrm{C}}$		$\log_{10} Z_{\mathrm{Ab}}$		$\log_{10} Z_{\mathrm{Env}}$	
100f	UB	LB	UB	LB	UB	LB	UB	LB
W	12.835	12.806	47.396	47.367	22.427	22.427	12.135	12.134
M	12.558	12.431	45.695	45.569	21.003	21.003	12.135	12.134
Q	12.495	12.370	44.582	44.458	19.953	19.953	12.135	12.134
Y	12.468	12.395	45.321	45.249	20.719	20.719	12.135	12.134
F	12.363	12.336	46.042	46.015	21.544	21.544	12.135	12.134
H_ϵ	12.355	12.303	45.848	45.798	21.359	21.359	12.135	12.134
N[1]	12.227	12.153	45.016	44.942	20.654	20.654	12.135	12.134
H_δ	12.159	12.101	45.910	45.853	21.617	21.617	12.135	12.134
R	11.232	10.980	45.929	45.686	22.571	22.563	12.135	12.134
K	10.778	10.496	45.163	44.884	22.254	22.250	12.135	12.134

[1] Wild-type amino-acid.

Table S4: Upper and lower bounds on the K^* score and partition functions for PG9 designs at residue 100k predicted using OSPREY, related to Results. H_δ and H_ε refer to the δ and ε protonation states of histidine.

	$\log_{10} K^*$		$\log_{10} Z_{\mathrm{C}}$		$\log_{10} Z_{\mathrm{Ab}}$		$\log_{10} Z_{\mathrm{Env}}$	
100k	UB	LB	UB	LB	UB	LB	UB	LB
E	10.630	10.318	35.980	35.780	4.009	4.009	21.453	21.340
D	8.719	8.401	34.576	34.371	4.517	4.517	21.453	21.340
W	8.346	8.026	34.930	34.723	5.244	5.244	21.453	21.340
Y[1]	7.738	7.415	34.413	34.203	5.336	5.335	21.453	21.340
Q	7.563	7.238	33.259	33.048	4.357	4.356	21.453	21.340
F	7.372	7.057	34.360	34.159	5.649	5.649	21.453	21.340
H_δ	7.285	6.955	33.941	33.725	5.317	5.316	21.453	21.340
M	7.213	6.879	33.630	33.409	5.077	5.077	21.453	21.340
H_ϵ	7.189	6.861	33.952	33.738	5.424	5.424	21.453	21.340
L	7.043	6.733	33.330	33.133	4.947	4.947	21.453	21.340
N	6.915	6.588	33.342	33.128	5.087	5.087	21.453	21.340
V	6.914	6.655	29.480	29.334	1.226	1.226	21.453	21.340
T	6.876	6.544	33.197	32.978	4.981	4.981	21.453	21.340
C	6.760	6.494	32.636	32.483	4.536	4.536	21.453	21.340
S	6.627	6.293	32.978	32.758	5.012	5.012	21.453	21.340
A	6.489	6.231	32.849	32.704	5.020	5.020	21.453	21.340
G	6.436	6.178	32.446	32.301	4.670	4.670	21.453	21.340
R	5.000	4.663	30.997	30.773	4.657	4.657	21.453	21.340
K	4.911	4.575	30.690	30.467	4.440	4.440	21.453	21.340

[1] Wild-type amino-acid.

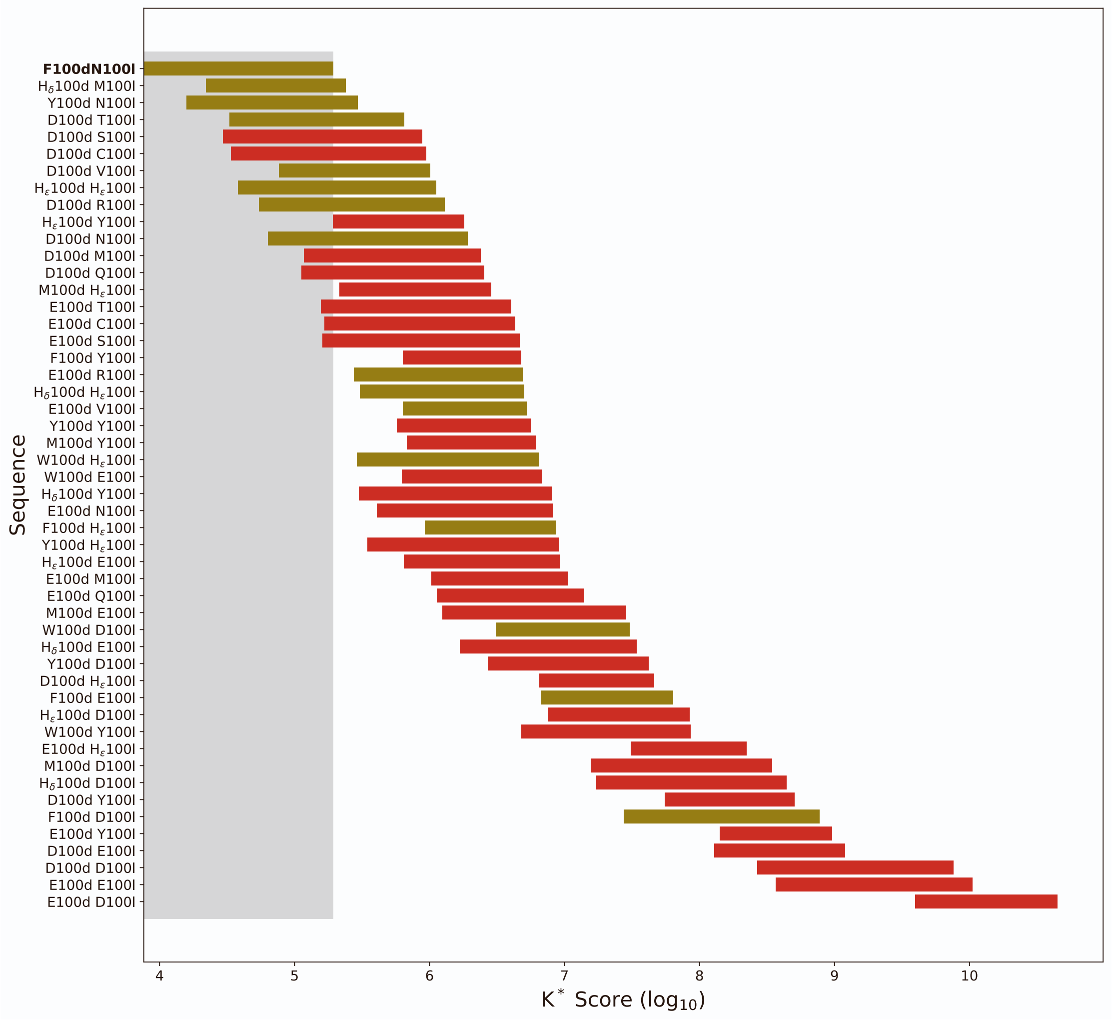
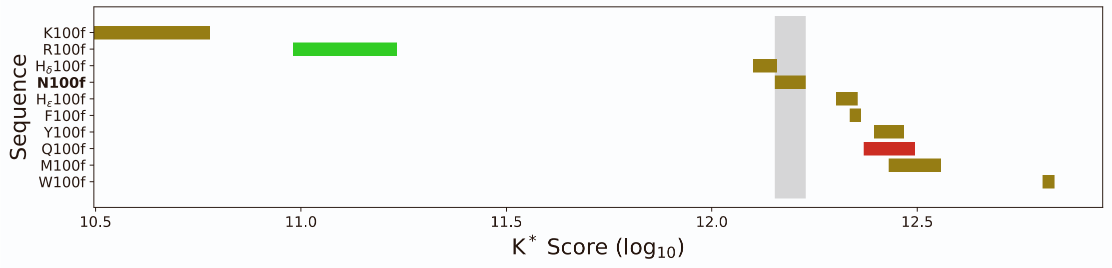
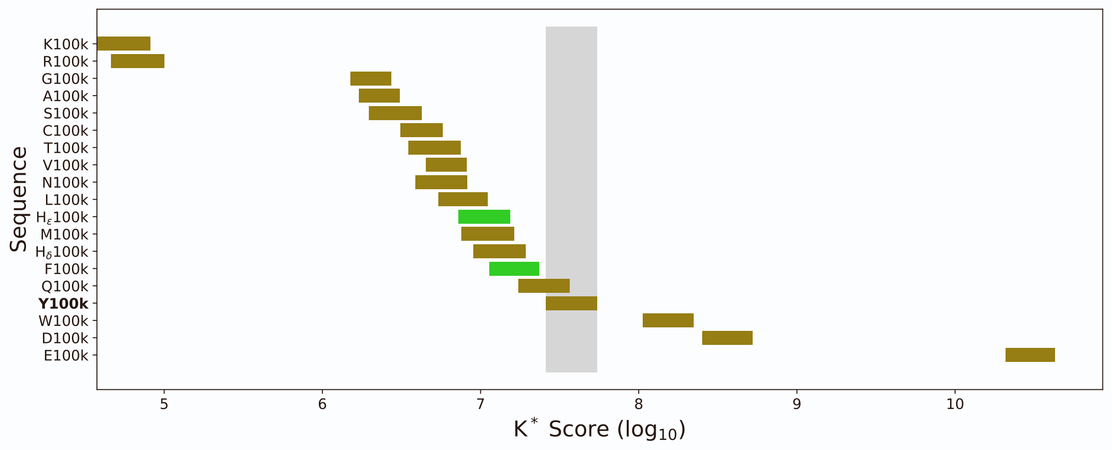


Figure S1: K^* scores for a double-mutation design of PGT145 at residues 100d and 100l, related to Results. Bounds on the K^* score for single mutations predicted using OSPREY are shown as horizontal bars. Pictured results are limited to the top 50 design predictions. Designs for which the unbound antibody is predicted to be more stable, approximately equally-stable, or less stable than wild-type are indicated by green, tan, or red coloring, respectively. Relative stability was estimated using the lower-bound on the partition-function value for the unbound antibody state. Wild-type antibody residue labeled in bold. The bounds on the K^* score for the wild-type design are indicated by the gray, shaded box.

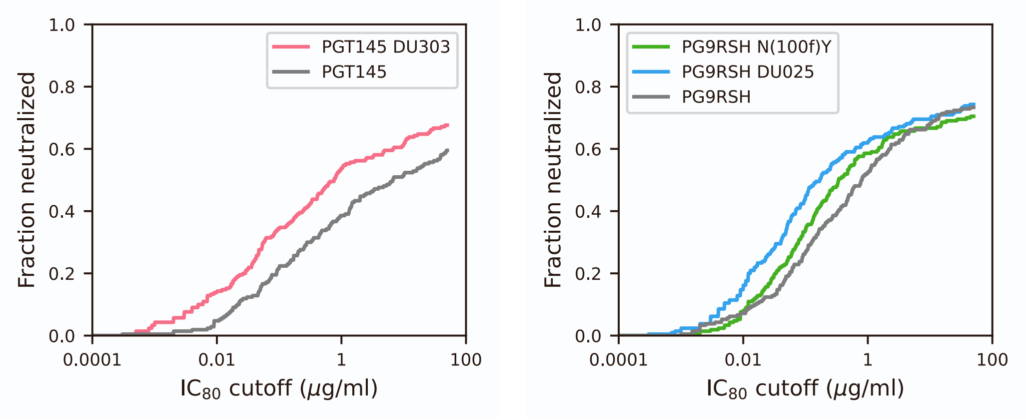


(A) K* scores for point mutation designs at PG9RSH residue 100f.

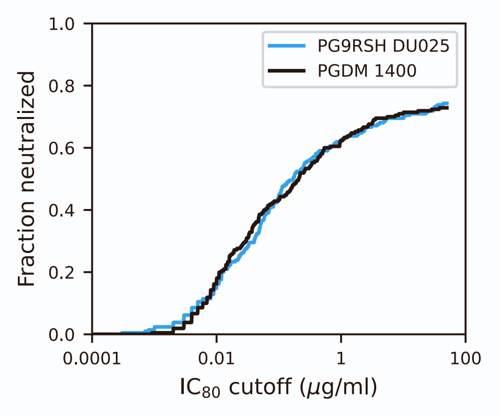


(B) K* scores for point mutation designs at PG9RSH residue 100k.

Figure S2: Selected design results for PG9RSH, related to Results. Bounds on the K* score for single mutations predicted using OSPREY are shown as horizontal bars. Designs for which the unbound antibody is predicted to be more stable, approximately equally-stable, or less stable than wild-type are indicated by green, tan, or red coloring, respectively. Relative stability was estimated using the lower-bound on the partition-function value for the unbound antibody state. Wild-type antibody residue labeled in bold. The bounds on the K* score for the wild-type design are indicated by the gray, shaded box.



(A) PGT145 **(B)** PG9RSH N(100f)Y and DU025



(C) DU025 and PGDM 1400

Figure S3: Variant neutralization improves over wild-type measured by IC_{80}, related to Figure 1. A,B) Breadth / potency curves for PGT145 and PG9RSH variants and controls, respectively. Curves represent the fraction of pseudoviruses that were neutralized with IC_{80} smaller than the given cutoff. An increase in breadth and potency is indicated by a shift upward and left. PGT145 variant DU303 and PG9RSH variant N(100f)Y and DU025 improve breadth and potency relative to wild-type. C) Breadth / potency curves for PG9RSH DU025 and PGDM1400. Despite slightly weaker median and mean neutralization potency, DU025 exhibits comparable breadth and potency to the best-in-class PGDM1400 antibody.

Table S5: PGT145 variants selected for small-panel neutralization assay, related to Results.

Identifier	Variant Mutations
DU301	F(100d)E
DU302	L(100e)E
DU303	N(100l)D
DU304	Y(100m)E
DU305	L(100e)E N(100l)D
DU306	L(100e)E N(100l)D Y(100m)M
DU307	F(100d)H L(100e)E N(100l)D
DU308	F(100d)E L(100e)E N(100l)R
DU309	F(100d)E L(100e)E N(100l)D
DU310	F(100d)H L(100e)E N(100l)D Y(100m)H

Table S6: PG9RSH variants selected for small-panel neutralization assay, related to Results.

Identifier	Variant Mutations
DU001	N(100c)H
DU002	N(100c)K
DU003	N(100c)M
DU004	N(100c)R
DU005	D100M R(100b)K N(100c)E
DU006	D100W R(100b)K N(100c)E
DU007	N(100f)K
DU008	N(100f)R
DU009	N(100f)W
DU010	N(100j)H
DU011	N(100f)Y
DU012	N(100f)W Y(100k)W
DU013	N(100f)Y Y(100k)W
DU014	N(100c)W N(100f)Y
DU015	N(100c)W N(100f)W
DU016	N(100c)F N(100f)Y
DU017	N(100c)F N(100f)W
DU018	D100E N(100c)Y N(100f)W
DU019	D100E N(100c)Y N(100f)Y
DU020	D100E R(100b)K N(100c)W N(100f)W
DU021	N(100f)M
DU022	N(100f)Q
DU023	F(100j)E
DU024	Y(100k)E
DU025	Y(100k)D
DU026	N(100c)F N(100f)W F(100j)E
DU027	N(100c)F N(100f)W Y(100k)E
DU028	N(100c)F N(100f)W Y(100k)D
DU029	N(100c)F Y(100e)D N(100f)W
DU030	N(100c)F Y(100e)E N(100f)W
DU031	N(100f)Y F(100j)E
DU032	N(100f)Y Y(100k)E
DU033	N(100f)Y Y(100k)D
DU034	Y(100e)D N(100f)Y

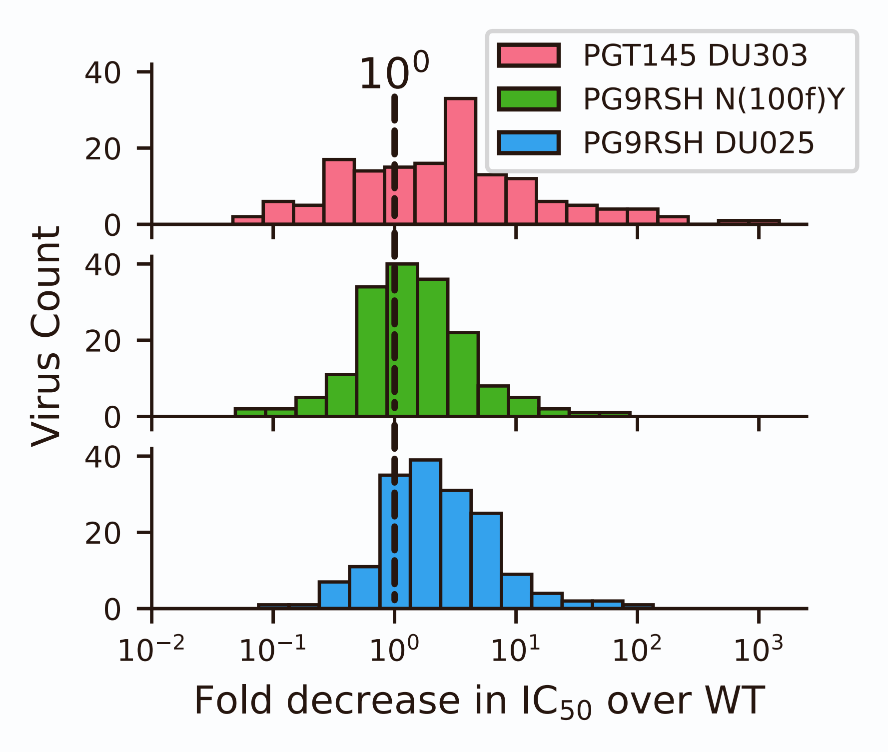


Figure S4: Fold change in neutralization for PGT145 and PG9RSH variants, related to Results. C) Histogram of fold decrease in IC_{50} values for each antibody / pseudovirus pair. Pairs where IC_{50} exceeded 50 μg/ml not shown. Dotted vertical line indicates a fold decrease of 1, or no change. All variants exhibit shifts to the right, indicating a trend of increased potency and breadth. This improvement is most marked in PGT145 DU303.

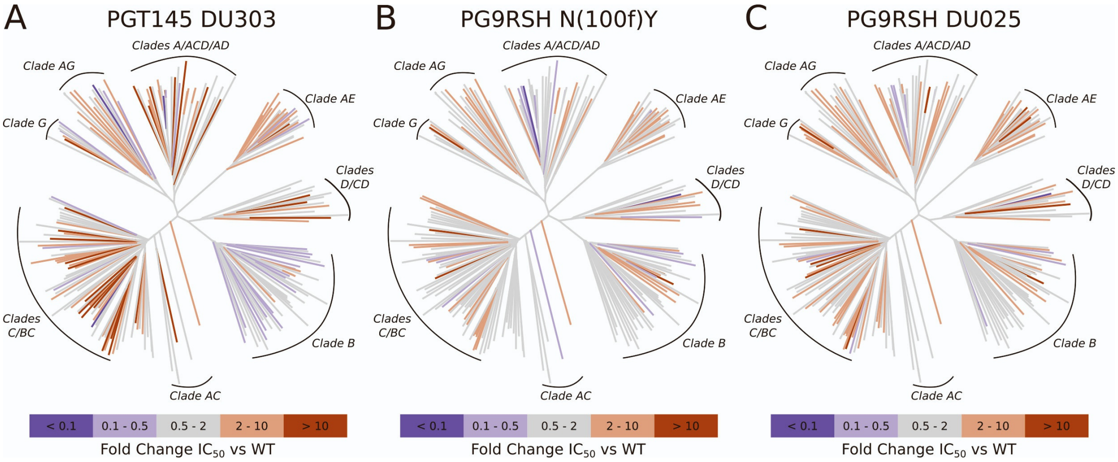


Figure S5: Difference dendrograms show patterns of neutralization relative to wild-type, related to Results. The differences in neutralization (fold-decrease in IC_{50}) between each antibody and its wild-type ancestor are shown using *difference dendrograms* for all large-panel viruses for variants DU303 (A), PG9RSH N(100f)Y (B), and DU025 (C), and respectively. Pseudoviruses are grouped into clades by sequence similarity, forming a tree graph. Internal branches in the tree, which denote groups of viruses, are colored gray. Terminal branches, corresponding to a single pseudovirus, are colored by the fold-decrease in IC_{50} relative to the wild-type antibody, where fold-increases in neutralization potency are indicated by a gradient of gray ($<$ 2-fold) to orange (2-10 fold) to dark red ($>$ 10-fold), and decreases in potency by a gradient of light purple (0.1-0.5 fold) to dark purple ($<$ 0.1-fold). A) DU303 exhibits increased neutralization for many viruses in clades C/BC, AG, A/ACD/AD, AE, and D/CD, but shows decreases in neutralization for viruses in clade B. B) PG9RSH N(100f)Y increases virus neutralization for almost all clades, with the possible exception of clade A/ACD/AD. Both the number of viruses with improved neutralization and the magnitude of changes in neutralization are less than for DU303. C) DU025 exhibits increased neutralization for some viruses in all clades, although the magnitude of these changes is less than for DU303.

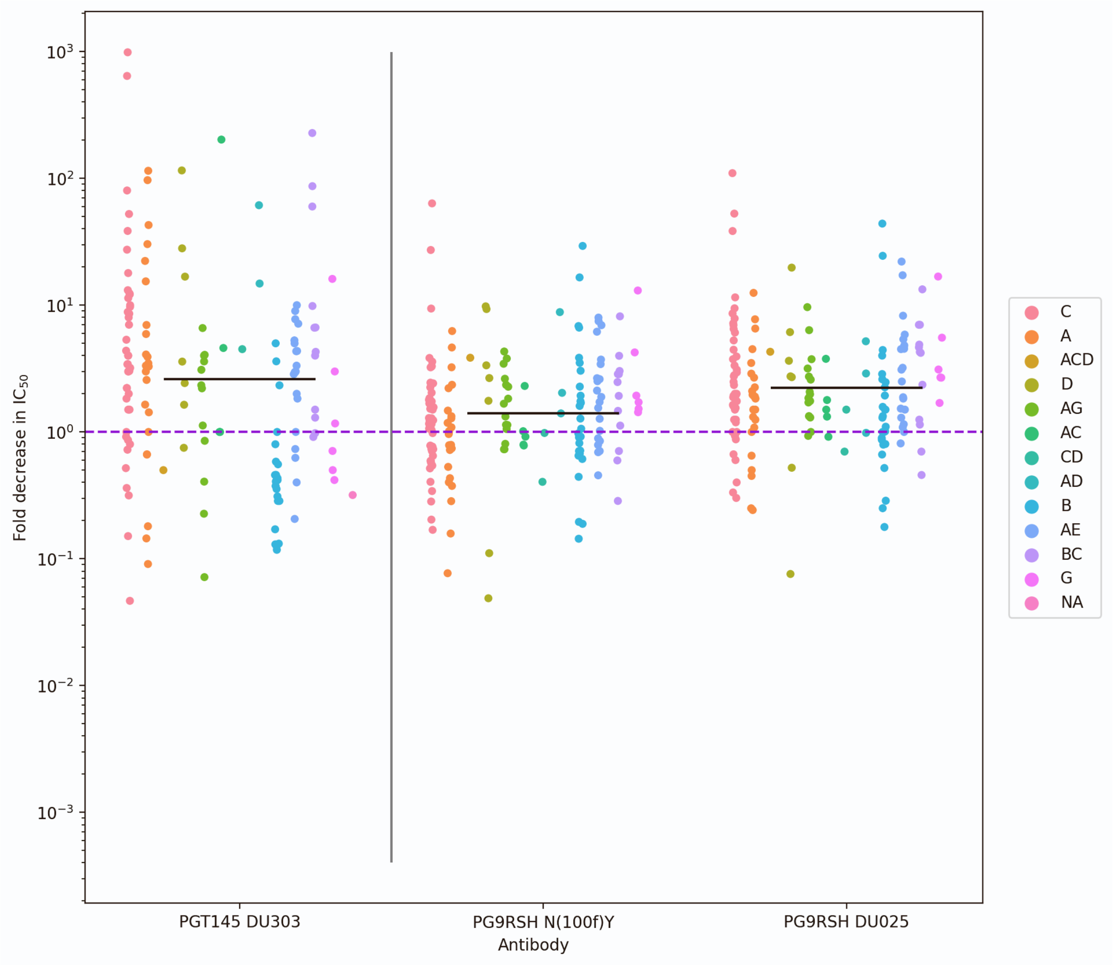
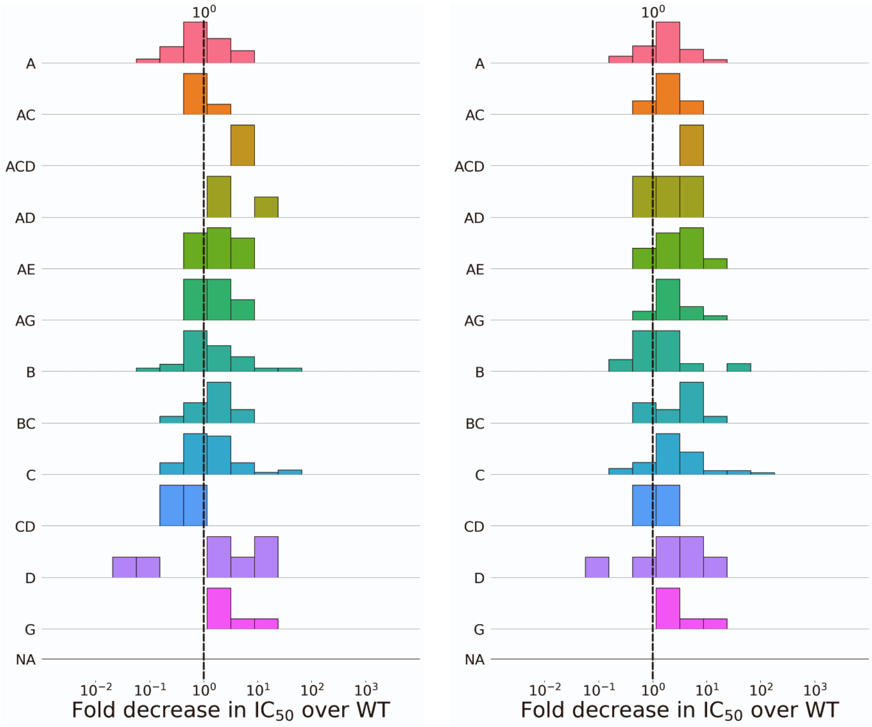
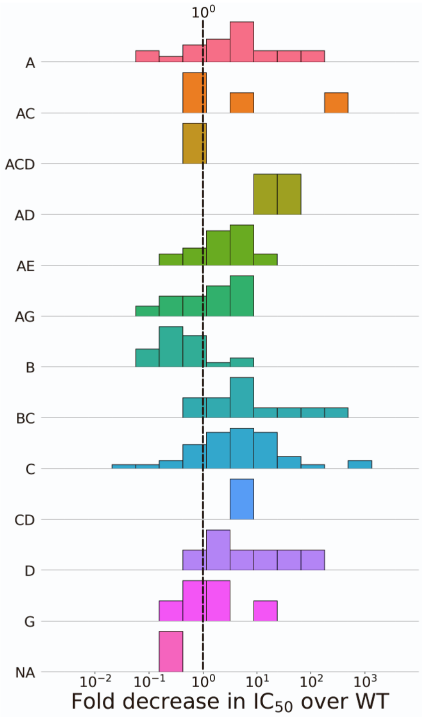


Figure S6: Fold decrease in IC_{50} over wild-type by pseudovirus, related to Results. Data from large neutralization panel excluding pseudoviruses for which antibody variant or wild-type $IC_{50} > 50\ \mu g/mL$. Colored by pseudovirus clade. Purple dotted line indicates a fold-change of 1, or no change. Horizontal black lines indicate the geometric mean for each variant.



(A) PG9RSH N(100f)Y **(B)** DU25



(C) DU303

Figure S7: Fold decrease IC50 by clade, histogram. Dotted vertical line indicates no change, related to Results. Rightward distribution shift indicates improvement.

Table S7: Summary of large-panel neutralization results (IC_{50}), related to Results.

		N(100f)Y[1,3]	DU025	PG9RSH	DU303	PGT145	PGDM1400
	# VS Assayed	208	208	208	208	208	208
# VS Neut.	IC_{50} 50 μg/mL	173	181	169	164	157	167
	IC_{50} <10 μg/mL	165	165	168	153	145	164
	IC_{50} <1.0 μg/mL	156	153	143	137	114	155
	IC_{50} <0.1 μg/mL	125	128	104	113	91	124
	IC_{50} <0.01 μg/mL	52	66	48	61	43	68
% VS Neut.	IC_{50} <50 μg/mL	83	87	81	79	75	80
	IC_{50} <10 μg/mL	79	79	81	74	70	79
	IC_{50} <1.0 μg/mL	75	74	69	66	55	75
	IC_{50} <0.1 μg/mL	60	62	50	54	44	60
	IC_{50} <0.01 μg/mL	25	32	23	29	21	33
	Median IC_{50} [2]	0.025	0.017	0.047	0.024	0.053	0.014
	Mean IC_{50} [1]	0.040	0.035	0.048	0.038	0.086	0.024

[1] Geometric mean calculated for samples with $IC_{50} < 50$ μg/mL.

[2] Median calculated for samples with $IC_{50} < 50$ μg/mL.

[3] Variant of PG9RSH.

Table S8: Summary of large-panel neutralization results (IC_{80}), related to Results.

		N(100f)Y[1,3]	DU025	PG9RSH	DU303	PGT145	PGDM1400
	# VS Assayed	208	208	208	208	208	208
# VS Neut.	IC_{80} 50 μg/mL	152	156	154	141	124	153
	IC_{80} <10 μg/mL	142	146	143	129	107	149
	IC_{80} <1.0 μg/mL	129	130	110	113	81	131
	IC_{80} <0.1 μg/mL	84	92	56	72	45	91
	IC_{80} <0.01 μg/mL	22	31	15	29	10	34
% VS Neut.	IC_{80} <50 μg/mL	73	75	74	68	60	74
	IC_{80} <10 μg/mL	69	70	69	62	51	72
	IC_{80} <1.0 μg/mL	62	63	53	54	39	63
	IC_{80} <0.1 μg/mL	40	44	27	35	22	44
	IC_{80} <0.01 μg/mL	11	15	7	14	5	16
	Median IC_{80} [2]	0.086	0.058	0.227	0.090	0.276	0.047
	Mean IC_{80} [1]	0.106	0.078	0.236	0.116	0.343	0.069

[1] Geometric mean calculated for samples with $IC_{80} < 50$ μg/mL.

[2] Median calculated for samples with $IC_{80} < 50$ μg/mL.

[3] Variant of PG9RSH.

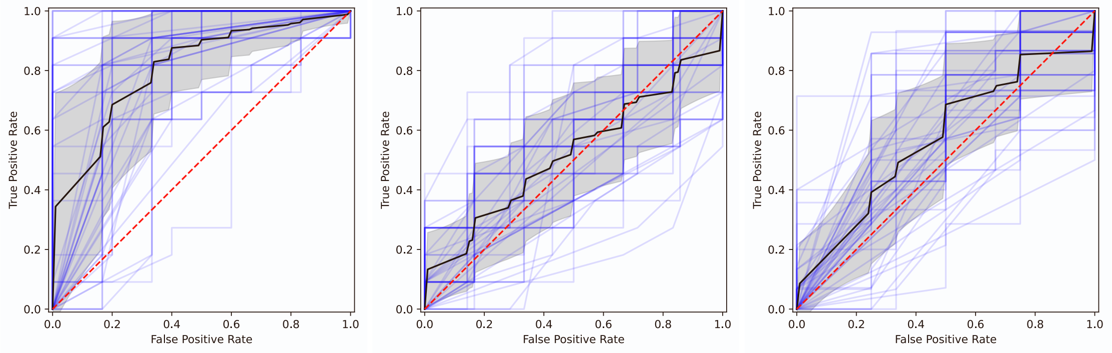
Table S9: Viruses with the largest differences in neutralization (IC_{50}) between wild-type (WT) and variant antibodies, related to Results.

Virus	Clade	Ab variant	WT IC_{50} [1]	Variant IC_{50} [1]	Fold-change
ZM233.6	C	PG9RSH DU025	0.11	0.001	110
Q259.17	A	PGT145 DU303	45.0	0.392	115
246-F3.C10.2	AC	PGT145 DU303	17.4	0.086	202
CH038.12	BC	PGT145 DU303	46.5	0.204	228
16936-2.21	C	PGT145 DU303	5.14	0.008	643
CAP256.206.C9	C	PGT145 DU303	3.95	0.004	988
191821.E6.1	D	PGT145 DU303	3.35	0.029	116

[1] IC_{50} values in μg/mL.

Table S10: Results of (5-times) repeated, nested 10-fold CV, related to Results. Inner 10-fold cross-validation was performed to tune hyperparameters, hyperparameter choices were evaluated by AUC.

bNAb model	**ROC AUC**	**Accuracy**	**F1 Score**
DU303	0.807 ± 0.128	0.732 ± 0.086	0.823 ± 0.054
PG9RSH N(100f)Y	0.521 ± 0.134	0.596 ± 0.085	0.721 ± 0.076
DU025	0.571 ± 0.134	0.767 ± 0.051	0.864 ± 0.038



(A) ROC curves for DU303 model. **(B)** ROC curves for PG9RSH N(100f)Y model. **(C)** ROC curves for PG9RSH N(100f)Y model.

Figure S8: ROC curves for repeated, nested cross-validation of models, related to Results. Models were evaluated using (5-times) repeated, 10-fold nested cross-validation. ROC curves for outer cross-validation steps are shown as blue traces. The mean ROC curve over all outer cross-validation steps is shown as a black trace, with shaded gray areas representing ± 1 standard deviation. Models for DU303 (A) give good predictive power on average, while models for PG9RSH N(100f)Y (B) and DU025 (C) fail to capture meaningful patterns.

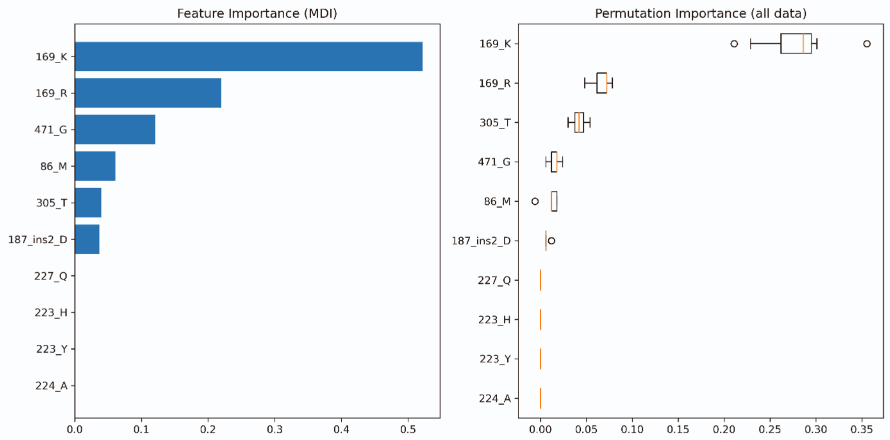


Figure S9: Variable importances for DU303 bNAb model, related to Results. Env residues K169 and R169 are important features for predicting change in neutralization between PGT145 and DU303, indicated by two separate variable importance measures. (A) Residues ranked by mean decrease in impurity (MDI). The MDI of a feature is the total reduction in Friedman mean squared error[1] for each tree-node that contains that feature, averaged over all decision trees in the ensemble. (B) Residues ranked by permutation importance (PI). The permutation importance measures the change in the loss function created by randomly permuting a feature[2]. This calculation was repeated 10 times for each feature. Replicates are visualized using box-and-whisker plots in which boxes extend from the first to third quartiles of the data, and whiskers extend by 1.5x the inter-quartile range.

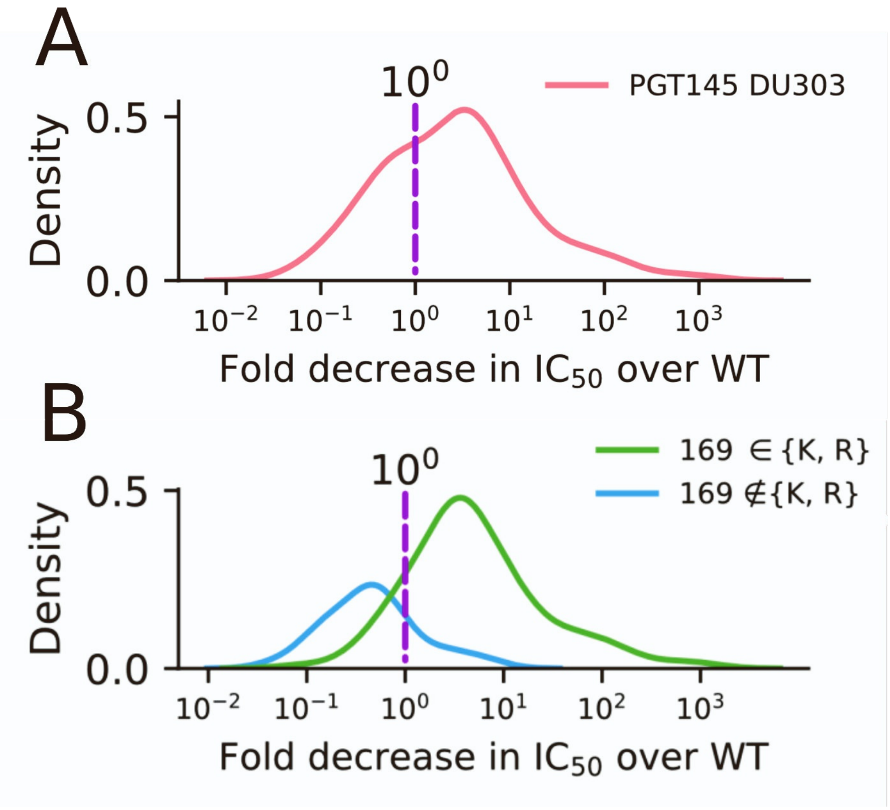


Figure S10: Distribution of changes in viral neutralization highlights the importance of Env residue 169 for neutralization by DU303, related to Results. Gaussian kernel density estimation was performed on data from large-panel neutralization assays to estimate probability distributions over the change in viral neutralization between wild-type PGT145 and variant DU303. Probability is shown on the y-axis, fold-decrease in IC_{50} shown on the x-axis. No change in neutralization (10^0 fold-decrease) is indicated with a purple, dashed vertical line. A) The distribution over all viruses with neutralization less than 50 μg/mL (pink) is shifted to the right of the vertical line, indicating that, in general, viruses are neutralized more potently by the DU303 variant than wild-type PGT145. However, the presence of a shoulder and long left distribution tail suggests that there is a significant population of viruses that are neutralized less-potently by DU303. B) Distributions over viruses with Env proteins that contain a positively-charged residue at position 169 (shown in green), or viruses without a positively-charged residue at Env position 169 (shown in blue) are well-separated. This indicates the importance of the amino-acid at this position to the neutralization by DU303. Predictive models (Figure S9) rely heavily on this correlation.

Table S11: Cryo-EM data collection and refinement statistics, related to STAR Methods.

	PG9RSH N(100f)Y BG505 DS-SOSIP.664	PG9RSH DU025 BG505 DS-SOSIP.664	PGT145 DU303 BG505 DS-SOSIP.664
EMDB ID	EMD-29248	EMD-29264	EMD-29288
PDB ID	8FK5	8FL1	8FLW
Microscope	FEI Titan Krios	FEI Titan Krios	FEI Titan Krios
Voltage (kV)	300	300	300
Electron dose ($e^-/\text{Å}^2$)	63.75	63.75	63.75
Detector	Gatan K3	Gatan K3	Gatan K3
Pixel size (Å)	1.083	1.083	1.083
Defocus range (μm)	-0.8 to -2.5	-0.8 to -2.5	-0.8 to -2.5
Magnification	81,000	81,000	81,000
Software	cryoSparc V3.1	cryoSparc V3.1	cryoSparc V3.1
Particles	230,180	104,665	107,753
Symmetry	C1	C1	C1
Box size (pix)	340	340	340
Resolution (Å) ($\text{FSC}_{0.143}$)	3.40	3.75	3.58
Software	Phenix 1.19	Phenix 1.19	Phenix 1.19
Protein residues	1,975	1,972	1,971
Chimera CC	78	79	76
EMRinger score	2.67	2.27	3.01
R.M.S. deviations			
Bond lengths (Å)	0.003	0.002	0.002
Bond angles (°)	0.469	0.460	0.538
Molprobity score	1.43	1.28	1.45
Clash score	3.76	3.90	3.37
Favored rotamers (%)	99.77	100	99.94
Ramachandran			
Favored regions (%)	96.12	97.47	95.33
Disallowed regions (%)	0.05	0.00	0.00

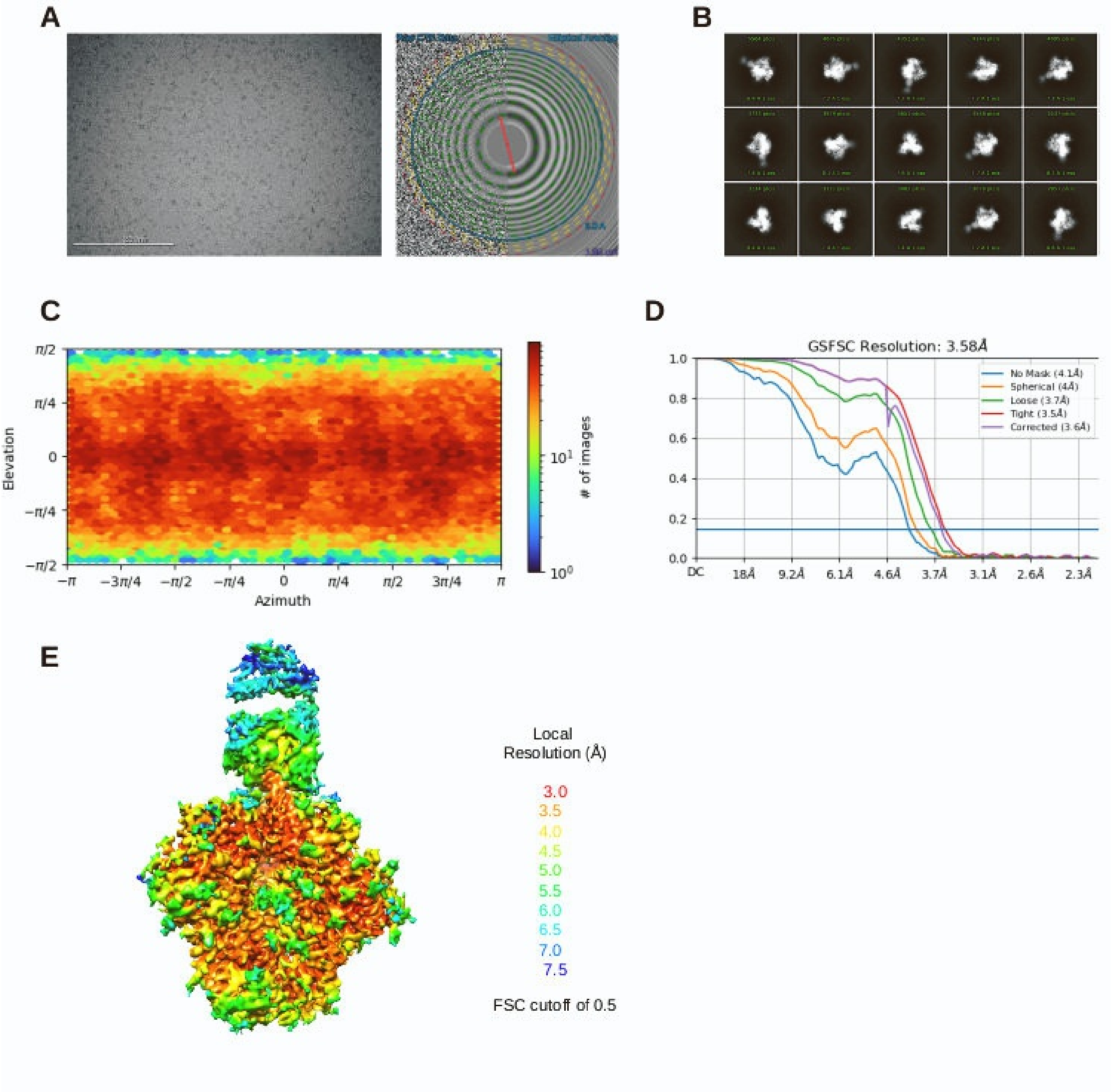


Figure S11: Cryo-EM Details of PG9RSH DU303 in complex with BG505 DS-SOSIP.664, related to Results. A) Representative micrograph and contrast transfer function (CTF) of micrograph. B) Representative 2D class averages. C) Heatmap of orientations of all particles used in the final refinement. D) The gold-standard Fourier shell correlation (FSC) results in a resolution of 3.58 Å using non-uniform refinement with C3 symmetry. E) The local resolution of the full map generated with cryoSPARC using an FSC of 0.5.

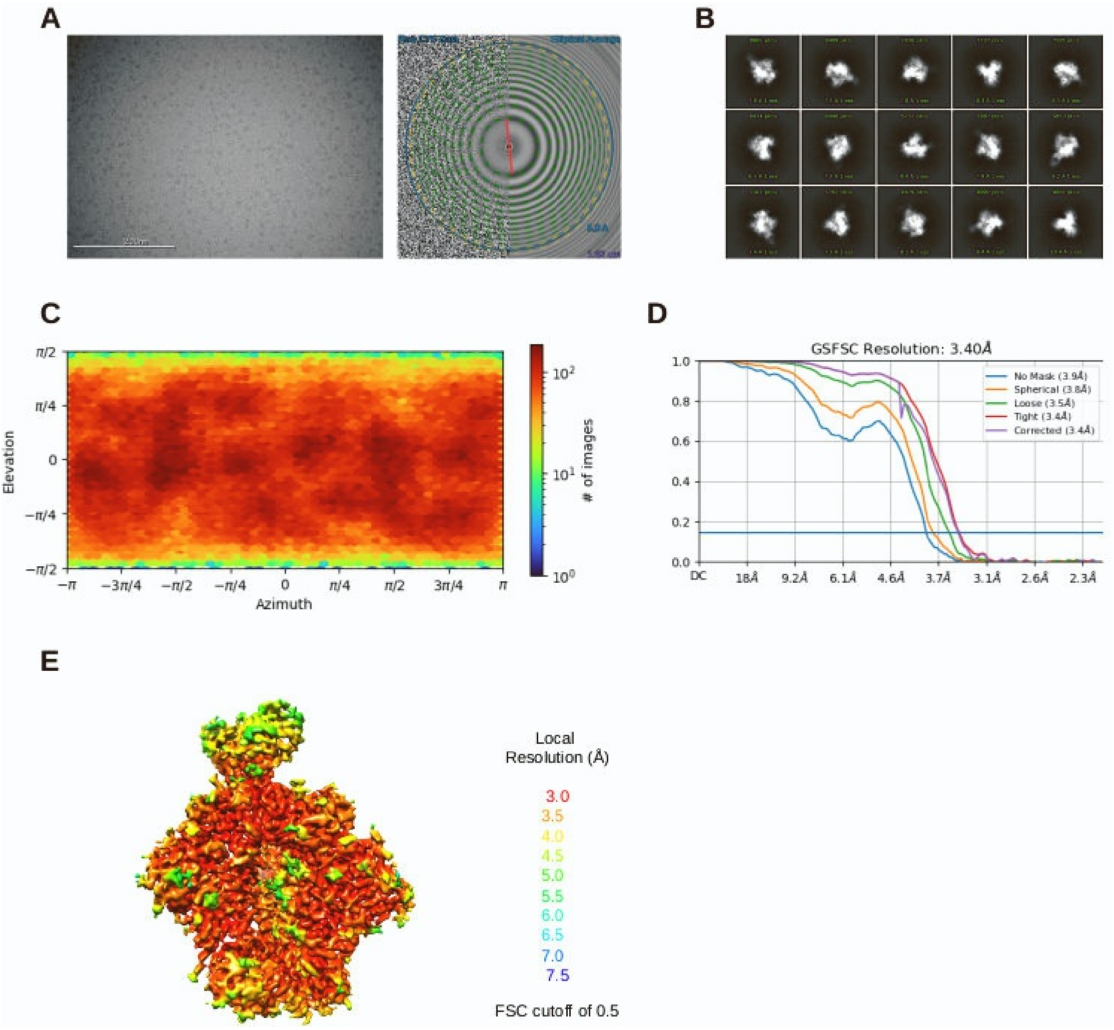


Figure S12: Cryo-EM Details of PG9RSH N(100f)Y in complex with BG505 DS-SOSIP.664, related to Results. A) Representative micrograph and contrast transfer function (CTF) of micrograph. B) Representative 2D class averages. C) Heatmap of orientations of all particles used in the final refinement. D) The gold-standard Fourier shell correlation (FSC) results in a resolution of 3.40 Å using non-uniform refinement with C3 symmetry. E) The local resolution of the full map generated with cryoSPARC using an FSC of 0.5.

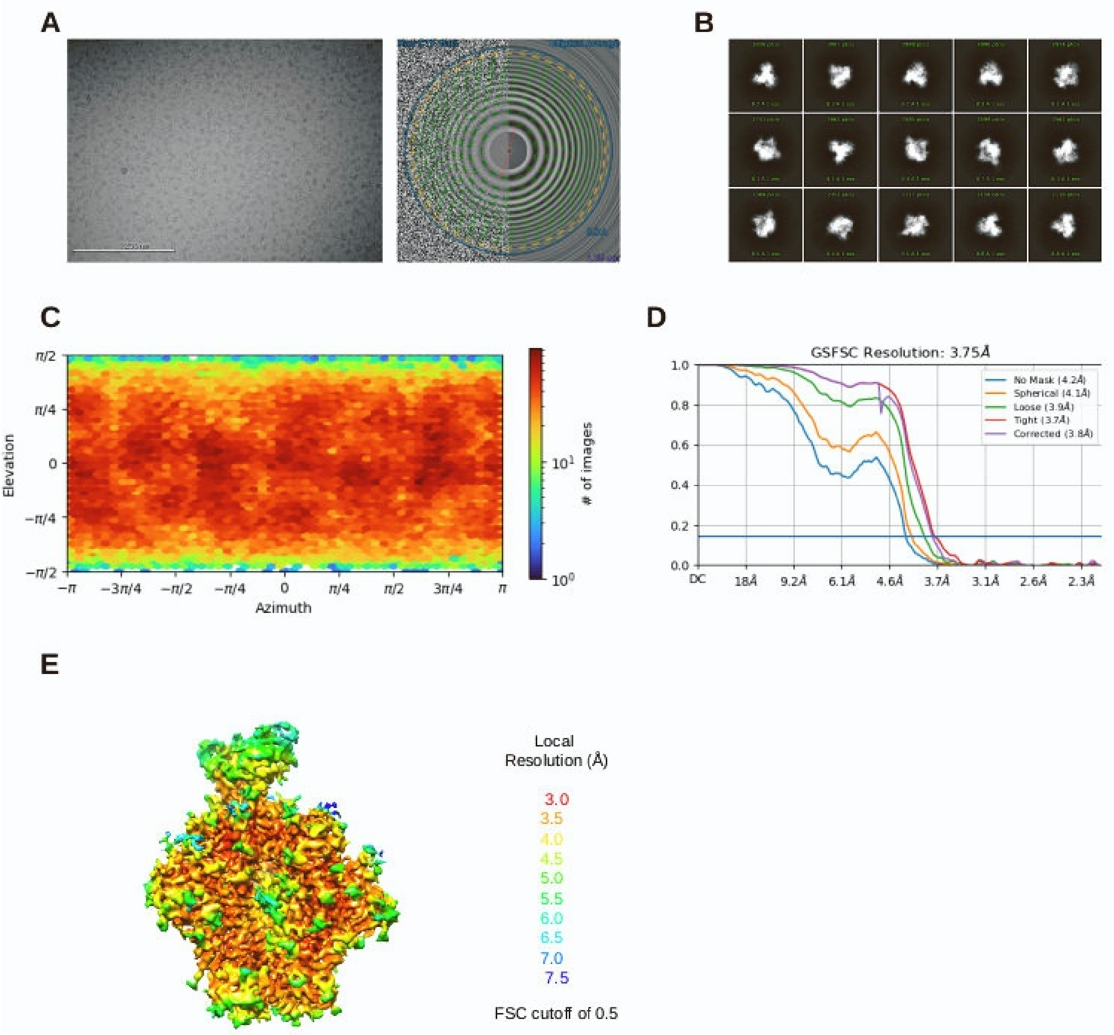
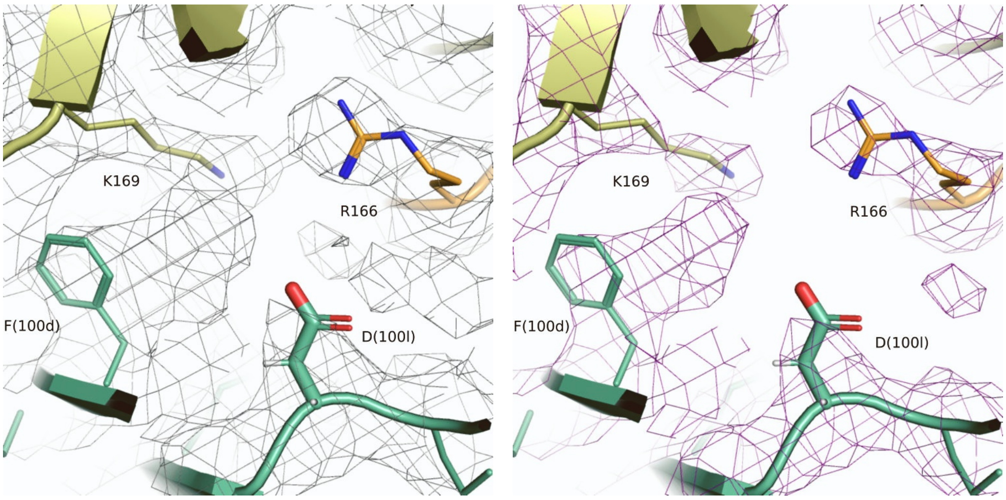
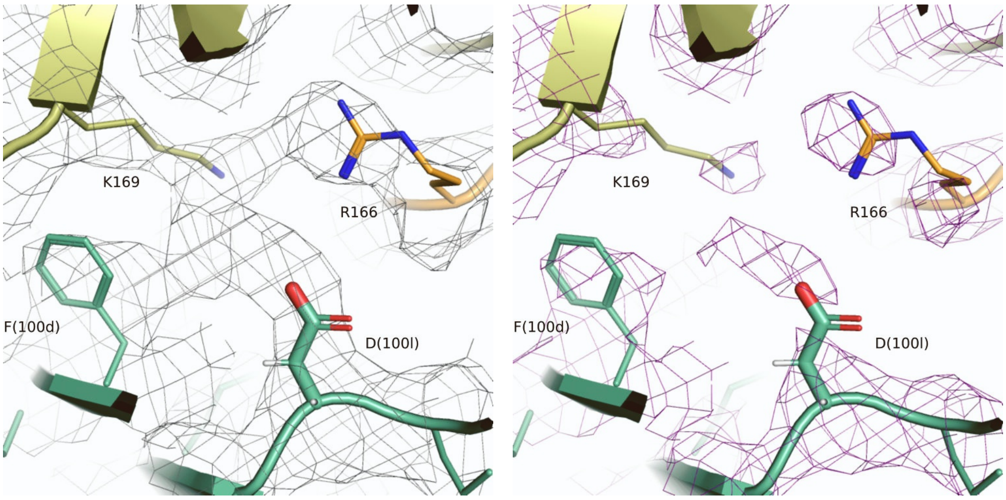


Figure S13: Cryo-EM Details of PG9RSH DU025 in complex with BG505 DS-SOSIP.664, related to Results. A) Representative micrograph and contrast transfer function (CTF) of micrograph. B) Representative 2D class averages. C) Heatmap of orientations of all particles used in the final refinement. D) The gold-standard Fourier shell correlation (FSC) results in a resolution of 3.75 Å using non-uniform refinement with C3 symmetry. E) The local resolution of the full map generated with cryoSPARC using an FSC of 0.5.

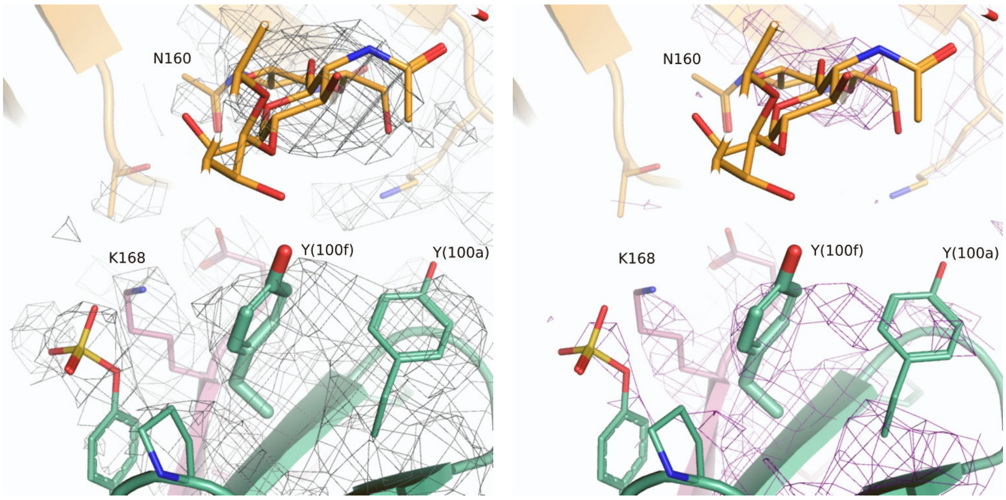


(A) Electron density isosurfaces (map 1) and atomic model around PGT145 DU303 residue D(100l).

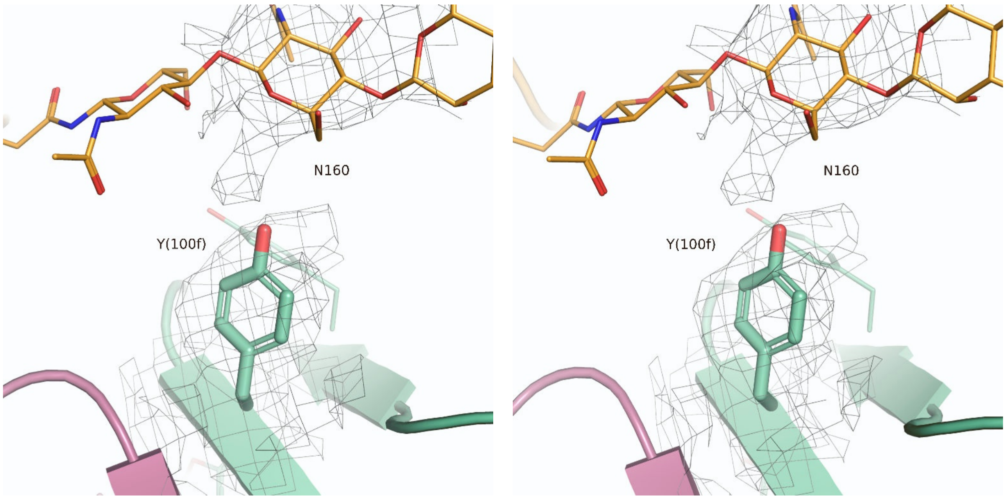


(B) Electron density isosurfaces (map 2) and atomic model around PGT145 DU303 residue D(100l).

Figure S14: Cryo-EM density near PGT145 DU303 residue D(100l), related to Results. Backbone shown in ribbon representation with glycans and amino acids shown as lines, and residue D(100l) shown as sticks. The interacting gp120 subunits are shown in orange and yellow, with the DU303 CDRH3 (residues 95-102) shown in green. The 1.2σ isosurface mesh shown in gray, the 3.0σ isosurface mesh shown in purple. Well-resolved side-chain density for gp120 residues K169 and R166 suggests potential electrostatic interactions with D(100l). A density bulge near residue F(100d) may indicate the presence of an alternative side-chain configuration that could interact with both D(100l) and gp120 residue K169.

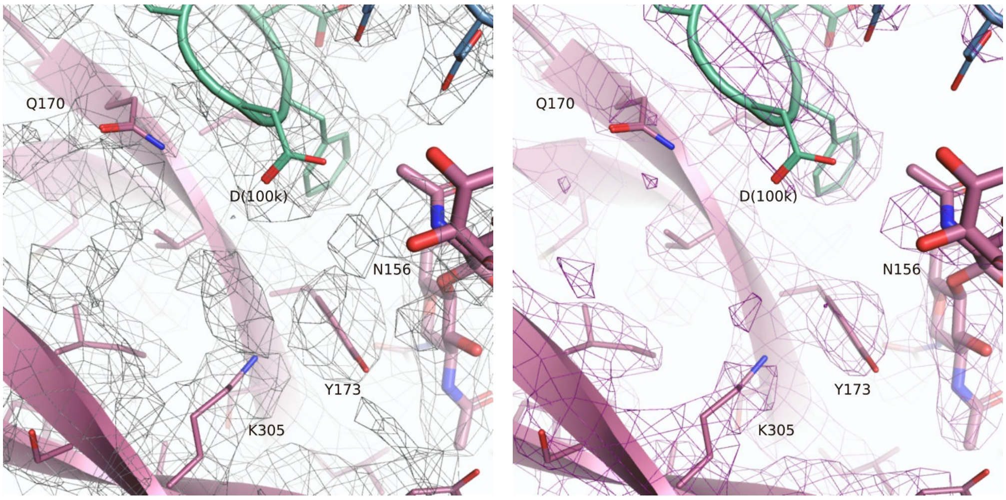


(A) Electron density isosurfaces and atomic model around PG9RSH N(100f)Y.

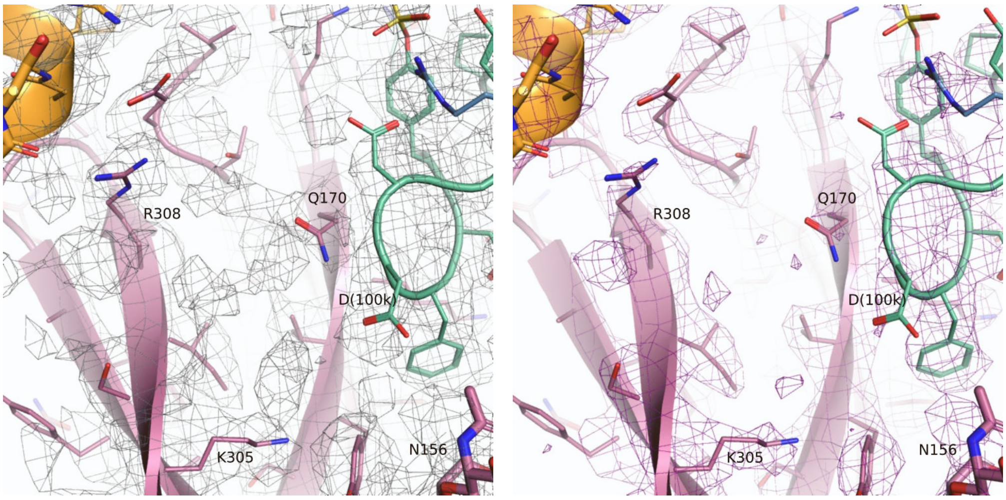


(B) Stereo view of low-density peaks around PG9RSH N(100f)Y.

Figure S15: Cryo-EM density near PG9RSH N(100f)Y, related to Results. Backbone shown in ribbon representation with glycans and amino acids shown as lines, and residue Y(100f) shown as sticks. The interacting gp120 subunits are shown in red and orange, with the PG9RSH N(100f)Y CDRH3 (residues 95-102) shown in green. (A) The 1.2σ isosurface mesh shown in gray, the 3.0σ isosurface mesh shown in purple. (B) The 0.5σ isosurface mesh shown in gray. Well-resolved side-chain density for gp120 residue K168 indicates that these side chains form favorable geometry for a pi-cation interaction. Density between residue Y(100f) and Y(100a) may indicate the presence of alternative side-chain conformations. A low-density peak between glycan N160 and Y(100f) indicates that this designed variant may form favorable interactions with the Env glycan shield.

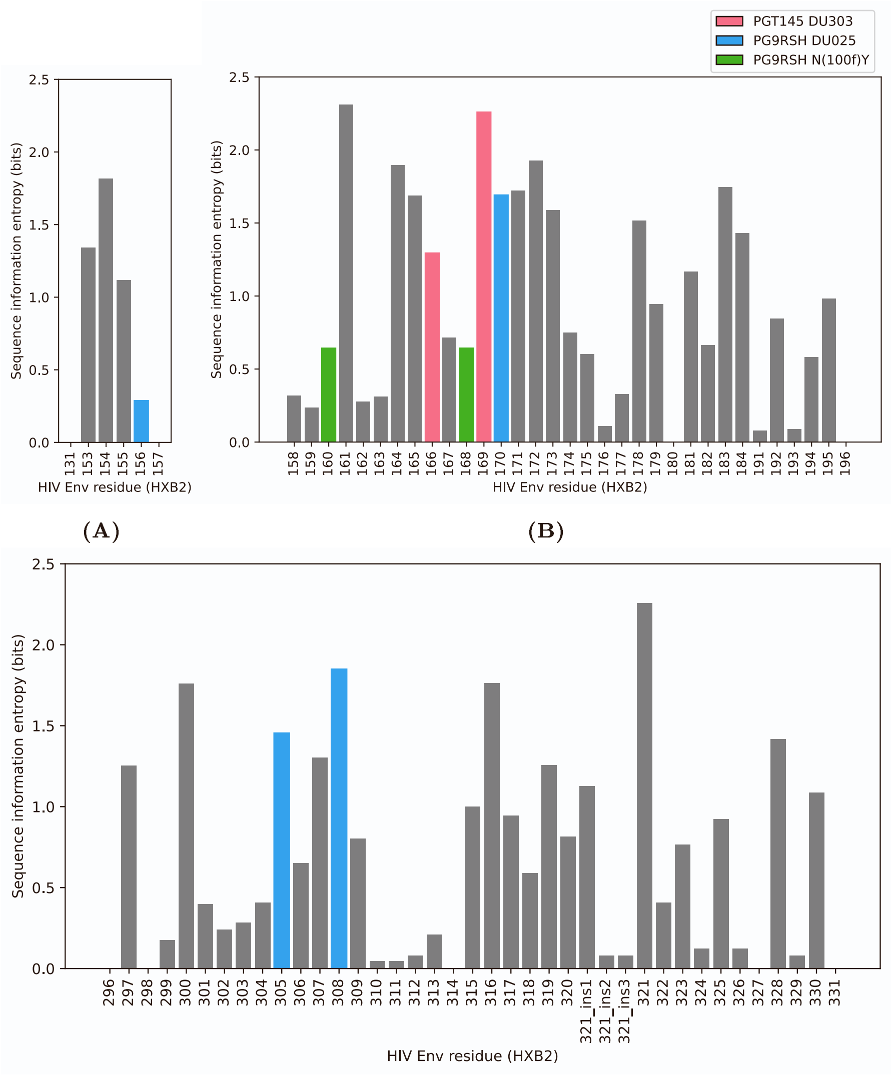


(A) Electron density isosurfaces and atomic model around PG9RSH DU025 residue D(100k).



(B) Electron density isosurfaces and atomic model around PG9RSH DU025 residue D(100k) and gp120 residue R308.

Figure S16: Cryo-EM density near PG9RSH DU025 residue D(100k), related to Results. Backbone shown in ribbon representation with glycans and amino acids shown as lines, and residue D(100l) shown as sticks. The interacting gp120 subunit is shown in red, with the DU303 CDRH3 (residues 95-102) shown in green and the remaining heavy-chain residues shown in blue. The 1.2σ isosurface mesh shown in gray, the 3.0σ isosurface mesh shown in purple. The D(100k) side-chain is well-resolved at 1.2σ, and retains a small but noticeable peak at 3.0σ. Side chains for gp120 residues Q170, K305, and Y173 are well-resolved at both levels, as are the first two GlcNAc monomers of glycan N156. Interestingly, several unassigned density peaks arise at 1.0σ located between the gp120 V2 and V3 loops. These could indicate solvent or transient glycan interactions, as only the core conserved N-linked glycan structure is modeled here. Density around gp120 residue R308 may indicate multiple populated rotamers.



(A)

(B)

(C)

Figure S17: Sequence information entropy of variable loops in large panel, related to Discussion. The number of bits of entropy for each residue in the V1, V2, and V3 loops (excluding hypervariable regions) of HIV Env in the large neutralization panel is shown in panels A, B, and C, respectively. Residue numbering and entropy calculations are based on alignment to the HXB2 reference sequence. Residues predicted to be important for neutralization of DU303, PG9RSH N(100f)Y, or DU025 by design ensembles or Cryo-EM structures are colored red, green, or blue, respectively.

References

[1] Jerome H. Friedman. "Greedy function approximation: A gradient boosting machine." In: *The Annals of Statistics* 29.5 (Oct. 2001), pp. 1189–1232. ISSN: 0090-5364. DOI: 10.1214/aos/1013203451. URL: https://projecteuclid.org/journals/annals-of-statistics/volume-29/issue-5/Greedy-function-approximation-A-gradient-boosting-machine/10.1214/aos/1013203451.full.

[2] Leo Breiman. *Random Forests*. Tech. rep. 2001, pp. 5–32.

Supplemental Results: Improved HIV-1 Neutralization Breadth and Potency of V2-Apex Antibodies by *In Silico* Design

Graham T. Holt[1,2,†], Jason Gorman[3,†], Siyu Wang[2], Anna U. Lowegard[1,2], Baoshan Zhang[3], Tracy Liu[3], Bob C. Lin[3], Mark K. Louder[3], Marcel S. Frenkel[4], Krisha McKee[3], Sijy O'Dell[3], Reda Rawi[3], Chen-Hsiang Shen[3], Nicole A. Doria-Rose[3], Peter D. Kwong[3,*], and Bruce R. Donald[1,4,5,6,*,‡]

[1]*Department of Computer Science, Duke University, Durham NC, USA*
[2]*Program in Computational Biology & Bioinformatics, Duke University, Durham NC, USA*
[3]*Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD, USA*
[4]*Department of Biochemistry, Duke University, Durham NC, USA*
[5]*Department of Mathematics, Duke University, Durham NC, USA*
[6]*Department of Chemistry, Duke University, Durham NC, USA*

Table SR1: Predicted K* scores for PGT145 WT and alanine mutants at residues of interest using OSPREY vs. corresponding IC_{50} data[1], related to Results.

	IC_{50} (μg/mL)[1]		K* score ($\log_{10}$K*)		Summary	
Residue	WT	Ala	WT	Ala	$-\Delta \log_{10} IC_{50}$	$\Delta \log_{10} K^*$
100d	0.010	2.380	3.438	2.483	−2.377	−0.955
100e	0.010	9.991	7.995	6.278	−3.000	−1.717
100l	0.010	0.004	15.454	15.155	0.398	−0.299
100m	0.010	>10	4.909	3.952	< -3	−0.957

[1] IC_{50} for BG505.Env.C2 pseudovirus.

Table SR2: Upper and lower bounds on the K* score and partition functions for PGT145 designs at residues 100e and 100m predicted using OSPREY, related to Results. H_δ and H_ε refer to the δ and ε protonation states of histidine.

		$\log_{10}$K*		$\log_{10} Z_{\mathrm{C}}$		$\log_{10} Z_{\mathrm{Ab}}$		$\log_{10} Z_{\mathrm{Env}}$	
100e	100m	UB	LB	UB	LB	UB	LB	UB	LB
E	E	18.239	16.797	29.182	28.687	6.963	6.468	4.971	4.476
E	D	18.202	16.802	29.035	28.540	6.862	6.367	4.961	4.466
E	Y	16.646	15.292	31.986	31.491	11.369	10.874	4.961	4.466
E	M	16.412	15.047	30.716[1]	30.221	10.213[1]	9.825	4.961[1]	4.479
E	H_ϵ	16.341	14.876	30.823	30.328	10.511	10.016	4.961	4.466
D	E	16.336	14.908	27.329	26.834	7.028	6.533	4.955	4.460
E	H_δ	16.327	14.900	30.937[1]	30.442	10.581[1]	10.143	4.961[1]	4.468
M	E	16.221	14.766	30.752[1]	30.257	10.536[1]	10.061	4.955[1]	4.470
D	D	16.161	14.718	27.003	26.508	6.870	6.375	4.961	4.466
E	Q	16.159	14.718	29.664	29.169	9.533	9.039	4.961	4.466
H_ϵ	E	16.148	14.671	30.558[1]	30.063	10.430[1]	9.941	4.962[1]	4.469
M	D	16.088	14.608	30.438[1]	29.943	10.372[1]	9.881	4.963[1]	4.469
H_ϵ	D	16.077	14.612	30.286[1]	29.792	10.222[1]	9.745	4.957[1]	4.465
L	E	15.976	14.515	30.573	30.078	10.627	10.132	4.960	4.465
Q	D	15.948	14.496	29.749[1]	29.254	9.801[1]	9.331	4.957[1]	4.470
Q	E	15.902	14.428	29.884[1]	29.389	9.997[1]	9.511	4.964[1]	4.471
L	D	15.896	14.467	30.340	29.845	10.478	9.983	4.956	4.461
E	I	15.824	14.569	30.803	30.308	10.999	10.505	4.969	4.474
E	T	15.689	14.222	29.558	29.063	9.897	9.403	4.961	4.466
E	V	15.647	14.217	29.620	29.126	10.003	9.508	4.961	4.466
E	S	15.525	14.056	29.006	28.511	9.514	9.019	4.958	4.463
E	C	15.510	14.096	28.813	28.318	9.332	8.837	4.961	4.466
E	A	15.467	14.041	28.794	28.299	9.355	8.860	4.961	4.466
E	G	15.352	13.938	28.250	27.755	8.926	8.431	4.961	4.466
H_δ	E	15.264	13.854	30.129[1]	29.634	10.821[1]	10.398	4.960[1]	4.468
H_δ	D	15.247	13.818	29.935[1]	29.441	10.660[1]	10.220	4.962[1]	4.469
T	E	15.090	13.609	29.368[1]	28.873	10.307[1]	9.812	4.958[1]	4.466
T	D	15.075	13.596	29.173[1]	28.678	10.128[1]	9.635	4.954[1]	4.462
V	D	15.032	13.594	28.938	28.443	9.935	9.440	4.961	4.466
V	E	15.021	13.682	29.194	28.699	10.202	9.707	4.961	4.466
D	Y	14.333	12.893	29.594	29.099	11.290	10.795	4.961	4.466
M	Y	14.053	12.579	32.213[1]	31.718	14.178[1]	13.688	4.960[1]	4.471
M	M	14.040	12.563	31.158[1]	30.663	13.138[1]	12.644	4.963[1]	4.474
D	H_ϵ	13.952	12.489	28.437	27.942	10.503	10.008	4.972	4.477
D	M	13.918	12.448	28.261	27.766	10.368	9.873	4.965	4.470
M	H_ϵ	13.909	12.432	31.375[1]	30.880	13.494[1]	13.004	4.954[1]	4.462
D	H_δ	13.881	12.400	28.497	28.002	10.648	10.153	4.957	4.462
H_ϵ	M	13.852	12.372	30.886[1]	30.391	13.057[1]	12.563	4.961[1]	4.470
Q	Y	13.836	12.389	31.446[1]	30.951	13.605[1]	13.127	4.957[1]	4.483
Q	M	13.830	12.348	30.382[1]	29.887	12.577[1]	12.083	4.961[1]	4.469
H_ϵ	Y	13.821	12.372	31.884[1]	31.389	14.058[1]	13.578	4.958[1]	4.485

Upper and lower bounds on the K* score and partition functions for PGT145 designs at residues 100e and 100m predicted using OSPREY (continued).

		$\log_{10}\text{K}^*$		$\log_{10} Z_{\text{C}}$		$\log_{10} Z_{\text{Ab}}$		$\log_{10} Z_{\text{Env}}$	
100e	100m	UB	LB	UB	LB	UB	LB	UB	LB
D	Q	13.788	12.340	27.322	26.827	9.562	9.068	4.961	4.466
M	Q	13.768	12.293	30.170^1	29.676	12.420^1	11.934	4.962^1	4.469
M	H_δ	13.768	12.312	31.141^1	30.646	13.383^1	12.911	4.952^1	4.462
L^2	Y^2	13.763	12.296	32.039	31.544	14.309	13.814	4.956	4.461
L	M	13.732	12.275	31.009	30.514	13.306	12.811	4.961	4.466
H_ϵ	H_ϵ	13.710	12.260	31.091^1	30.596	13.382^1	12.915	4.954^1	4.466
Q	H_ϵ	13.688	12.245	30.610^1	30.115	12.915^1	12.442	4.955^1	4.480
H_ϵ	Q	13.628	12.180	29.958^1	29.463	12.322^1	11.862	4.961^1	4.467
Q	H_δ	13.600	12.131	30.410^1	29.915	12.828^1	12.342	4.957^1	4.469
Q	Q	13.571	12.102	29.412^1	28.917	11.862^1	11.377	4.953^1	4.463
L	H_ϵ	13.571	12.196	31.225	30.730	13.683	13.188	4.961	4.466
H_ϵ	H_δ	13.546	12.076	30.835^1	30.341	13.304^1	12.823	4.960^1	4.466
L	Q	13.503	12.037	30.019	29.525	12.551	12.056	4.955	4.460
L	H_δ	13.503	12.024	30.964	30.469	13.492	12.997	4.959	4.464
D	I	13.356	11.935	28.198	27.703	10.860	10.365	4.972	4.477
D	T	13.305	11.829	27.164	26.669	9.892	9.397	4.957	4.462
M	T	13.295	11.815	30.043^1	29.548	12.772^1	12.280	4.961^1	4.467
H_δ	M	13.204	11.777	30.605^1	30.110	13.378^1	12.941	4.956^1	4.461
M	I	13.177	11.728	30.990^1	30.495	13.807^1	13.347	4.960^1	4.465
H_ϵ	T	13.142	11.666	29.815^1	29.320	12.693^1	12.201	4.961^1	4.472
D	V	13.132	11.699	27.154	26.659	10.051	9.556	4.961	4.466
M	C	13.100	11.626	29.174^1	28.679	12.093^1	11.604	4.960^1	4.469
M	S	13.097	11.620	29.423^1	28.928	12.352^1	11.863	4.955^1	4.464
Q	T	13.084	11.614	29.303^1	28.808	12.235^1	11.751	4.959^1	4.468
H_δ	Y	13.081	11.641	31.472^1	30.977	14.373^1	13.915	4.963^1	4.475
D	C	13.075	11.648	26.330	25.835	9.283	8.788	4.961	4.466
H_δ	H_ϵ	13.054	11.586	30.820^1	30.325	13.779^1	13.300	4.960^1	4.465
D	S	13.040	11.588	26.588	26.093	9.573	9.078	4.965	4.470
Q	I	13.026	11.620	30.246^1	29.751	13.175^1	12.757	4.956^1	4.462
L	T	13.018	11.541	29.892	29.397	12.905	12.410	4.958	4.464
H_ϵ	C	12.998	11.528	28.930^1	28.436	11.951^1	11.465	4.957^1	4.467
M	V	12.986	11.543	29.841^1	29.346	12.844^1	12.390	4.958^1	4.464
T	M	12.969	11.497	29.849^1	29.354	12.898^1	12.408	4.959^1	4.472
M	A	12.964	11.499	29.149^1	28.654	12.187^1	11.708	4.968^1	4.477
T	Y	12.955	11.476	30.892^1	30.397	13.958^1	13.466	4.962^1	4.471
E	R	12.953	11.495	30.116	29.622	13.182	12.687	4.972	4.477
Q	C	12.951	11.474	28.451^1	27.956	11.527^1	11.038	4.955^1	4.462
H_ϵ	S	12.945	11.464	29.188^1	28.693	12.276^1	11.782	4.954^1	4.462
D	A	12.934	11.566	26.339	25.844	9.434	8.939	4.961	4.466
V	Y	12.925	11.468	30.651	30.157	13.755	13.260	4.961	4.466
H_δ	Q	12.888	11.412	29.553^1	29.058	12.687^1	12.198	4.960^1	4.466
L	I	12.885	11.475	30.824	30.329	13.964	13.469	4.965	4.470
H_ϵ	I	12.871	11.445	30.569^1	30.074	13.668^1	13.229	4.961^1	4.469

Upper and lower bounds on the K* score and partition functions for PGT145 designs at residues 100e and 100m predicted using OSPREY (continued).

		$\log_{10}\mathrm{K}^*$		$\log_{10} Z_{\mathrm{C}}$		$\log_{10} Z_{\mathrm{Ab}}$		$\log_{10} Z_{\mathrm{Env}}$	
100e	100m	UB	LB	UB	LB	UB	LB	UB	LB
Q	S	12.859	11.401	28.679[1]	28.184	11.820[1]	11.348	4.963[1]	4.472
D	G	12.834	11.444	25.793	25.298	8.987	8.492	4.961	4.466
L	C	12.830	11.370	29.041	28.546	12.232	11.737	4.969	4.474
L	S	12.824	11.351	29.274	28.779	12.478	11.984	4.961	4.466
M	G	12.815	11.402	28.560[1]	28.066	11.707[1]	11.280	4.957[1]	4.465
T	H_ϵ	12.785	11.313	30.043[1]	29.549	13.275[1]	12.781	4.960[1]	4.477
V	M	12.782	11.418	29.605	29.110	12.852	12.357	4.961	4.466
H_δ	H_δ	12.776	11.418	30.460[1]	29.965	13.587[1]	13.215	4.960[1]	4.469
H_ϵ	V	12.758	11.390	29.549[1]	29.054	12.707[1]	12.328	4.958[1]	4.464
Q	A	12.756	11.308	28.400[1]	27.905	11.638[1]	11.167	4.958[1]	4.477
Q	V	12.756	11.340	29.097[1]	28.602	12.304[1]	11.860	4.958[1]	4.481
V	H_ϵ	12.750	11.299	29.806	29.311	13.089	12.595	4.956	4.461
H_ϵ	A	12.673	11.362	28.899[1]	28.405	12.089[1]	11.762	4.954[1]	4.464
L	V	12.657	11.286	29.669	29.174	13.034	12.540	4.967	4.472
T	Q	12.634	11.162	28.834[1]	28.340	12.219[1]	11.728	4.959[1]	4.472
T	H_δ	12.623	11.176	29.808[1]	29.313	13.182[1]	12.707	4.954[1]	4.477
L	A	12.584	11.262	28.996	28.501	12.447	11.952	4.955	4.460
V	H_δ	12.582	11.158	29.556	29.061	13.002	12.507	4.961	4.466
Q	G	12.581	11.180	27.815[1]	27.320	11.180[1]	10.761	4.960[1]	4.472
V	Q	12.549	11.175	28.609	28.114	12.086	11.591	4.964	4.469
L	G	12.514	11.110	28.402	27.907	11.916	11.421	4.961	4.466
H_ϵ	G	12.473	11.001	28.303[1]	27.809	11.846[1]	11.362	4.962[1]	4.469
H_δ	T	12.406	10.939	29.446[1]	28.951	13.052[1]	12.573	4.961[1]	4.467
H_δ	I	12.405	10.934	30.324[1]	29.829	13.942[1]	13.454	4.953[1]	4.466
H_δ	S	12.223	10.742	28.828[1]	28.333	12.633[1]	12.140	4.959[1]	4.465
H_δ	C	12.208	10.735	28.573[1]	28.078	12.386[1]	11.897	4.957[1]	4.468
T	T	12.158	10.682	28.704[1]	28.209	12.570[1]	12.077	4.957[1]	4.469
V	T	12.128	10.664	28.467	27.972	12.368	11.873	4.961	4.466
T	I	12.119	10.644	29.643[1]	29.148	13.543[1]	13.056	4.961[1]	4.468
H_δ	G	12.047	10.568	28.103[1]	27.608	12.079[1]	11.590	4.960[1]	4.466
H_δ	A	12.023	10.651	28.534[1]	28.039	12.436[1]	12.053	4.952[1]	4.458
T	C	11.968	10.501	27.838[1]	27.344	11.882[1]	11.404	4.961[1]	4.467
T	S	11.953	10.471	28.090[1]	27.595	12.167[1]	11.674	4.957[1]	4.463
V	S	11.926	10.450	27.840	27.345	11.942	11.447	4.961	4.466
V	I	11.921	10.641	29.419	28.924	13.527	13.033	4.961	4.466
V	C	11.897	10.467	27.575	27.080	11.703	11.208	4.964	4.469
T	V	11.857	10.408	28.513[1]	28.018	12.649[1]	12.188	4.962[1]	4.468
T	A	11.835	10.372	27.819[1]	27.324	12.006[1]	11.513	4.946[1]	4.471
H_δ	V	11.835	10.536	29.093[1]	28.598	13.100[1]	12.789	4.962[1]	4.469
M	R	11.784	10.304	29.923[1]	29.428	14.160[1]	13.668	4.964[1]	4.471
V	V	11.739	10.375	28.272	27.777	12.568	12.073	4.955	4.460
T	G	11.704	10.248	27.234[1]	26.739	11.537[1]	11.046	4.954[1]	4.484
V	A	11.692	10.343	27.564	27.069	11.901	11.406	4.961	4.466

Upper and lower bounds on the K* score and partition functions for PGT145 designs at residues 100e and 100m predicted using OSPREY (continued).

		$\log_{10}$K*		$\log_{10} Z_{\mathrm{C}}$		$\log_{10} Z_{\mathrm{Ab}}$		$\log_{10} Z_{\mathrm{Env}}$	
100e	100m	UB	LB	UB	LB	UB	LB	UB	LB
Q	R	11.579	10.104	29.258[1]	28.763	13.705[1]	13.216	4.954[1]	4.463
L	R	11.473	10.010	29.745	29.250	14.301	13.806	4.961	4.466
V	G	11.456	10.215	26.956	26.461	11.533	11.038	4.957	4.462
H_ϵ	R	11.434	9.963	29.676[1]	29.181	14.263[1]	13.780	4.956[1]	4.463
H_δ	R	10.948	9.473	29.260[1]	28.765	14.333[1]	13.848	4.959[1]	4.464
V	R	10.608	9.132	28.326	27.832	13.748	13.253	4.961	4.466
T	R	10.593	9.118	28.603[1]	28.109	14.035[1]	13.540	4.956[1]	4.470
D	R	10.578	9.151	27.473	26.979	12.914	12.419	4.972	4.477

[1] Upper bounds computed using lower bound and $\epsilon = 0.68$.

[2] Wild-type amino-acid.

Table SR3: Root mean-squared deviation between existing structures of PGT145, related to Results.

	RMSD (Å)[1,2]					
PDB ID	8FLW[3]	6NIJ	58VL	8DVD	3U1S	5U1F
8FLW[3]	0.00 (140)					
6NIJ	1.95 (140)	0.00 (140)				
5V8L	1.32 (140)	1.69 (140)	0.00 (140)			
8DVD	1.40 (140)	1.86 (140)	1.03 (140)	0.00 (140)		
3U1S	1.35 (140)	1.66 (140)	0.76 (140)	0.76 (140)	0.00 (240)	
5U1F	1.37 (138)	1.69 (138)	0.80 (138)	0.84 (138)	1.16 (235)	0.00 (235)

[1] RMSD computed for all resolved PGT145 antibody heavy chain C_α atoms using the PyMOL `align` command with the `cycles=0` option.

[2] Numbers in parentheses represent the number of atoms included in the RMSD calculation.

[3] Cryo-EM structure of PGT145 DU303.

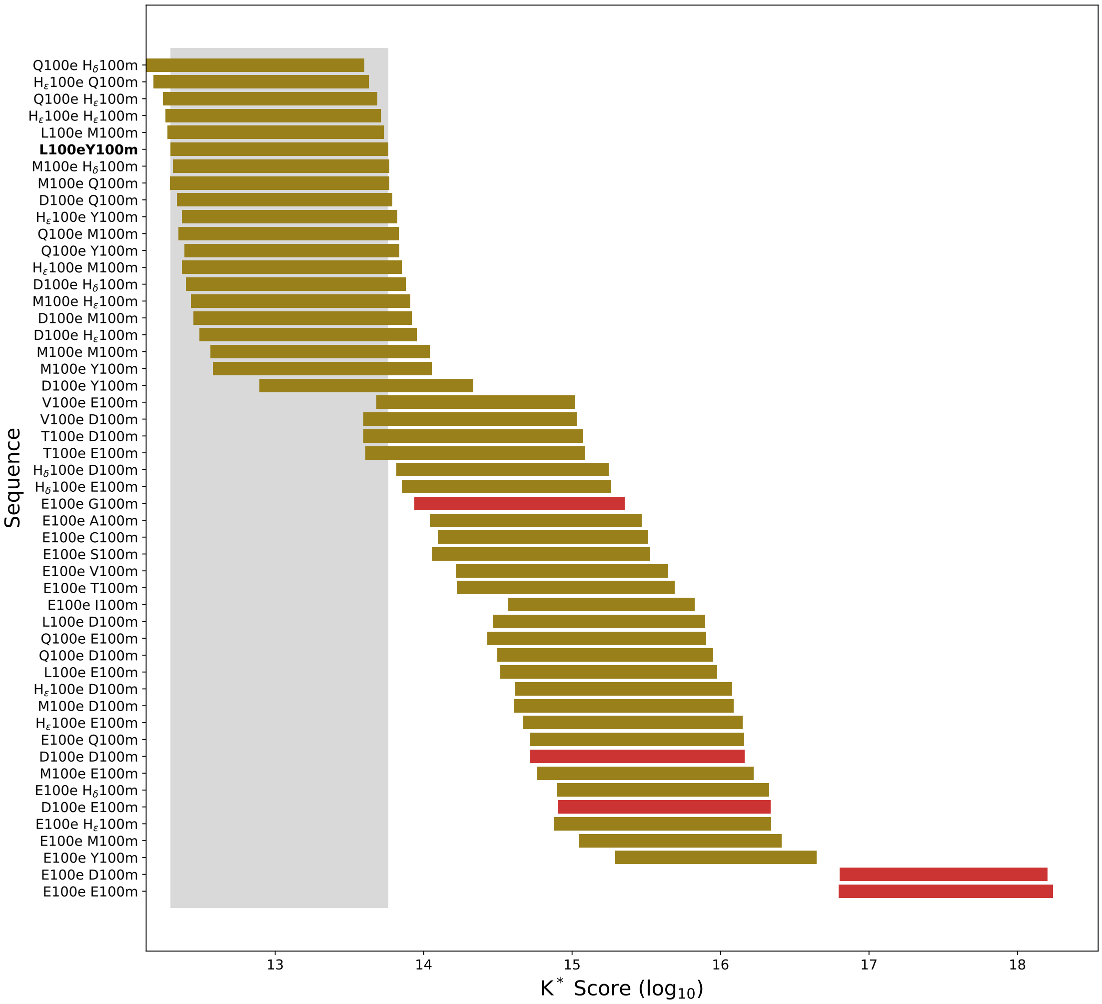


Figure SR1: K* scores for a double-mutation design of PGT145 at residues 100e and 100m, related to Results. Bounds on the K* score for single mutations predicted using OSPREY are shown as horizontal bars. Pictured results are limited to the top 50 design predictions. Designs for which the unbound antibody is predicted to be more stable, approximately equally-stable, or less stable than wild-type are indicated by green, tan, or red coloring, respectively. Relative stability was estimated using the lower-bound on the partition-function value for the unbound antibody state. Wild-type antibody residue labeled in bold. The bounds on the K* score for the wild-type design are indicated by the gray, shaded box.

Table SR4: Root mean-squared deviation between existing structures of PG9, related to Results.

	RMSD (Å)[1,2]							
PDB ID	8FK5[3]	8FL1[4]	7T77	7T76	3U4E	3U2S	4YAQ	5VJ6
8FK5[3]	0.00 (136)							
8FL1[4]	0.57 (136)	0.00 (136)						
7T77	0.61 (133)	0.85 (133)	0.00 (133)					
7T76	0.81 (135)	1.00 (135)	0.61 (133)	0.00 (136)				
3U4E	0.62 (136)	0.83 (136)	0.62 (133)	0.84 (135)	0.00 (239)			
3U2S	0.62 (136)	0.87 (136)	0.44 (133)	0.77 (135)	3.38 (230)	0.00 (230)		
4YAQ	0.69 (136)	0.92 (136)	0.53 (133)	0.85 (135)	5.14 (239)	1.94 (230)	0.00 (239)	
5VJ6	2.18 (135)	2.34 (135)	2.06 (133)	2.02 (135)	2.35 (229)	3.59 (229)	5.16 (229)	0.00 (229)

[1] RMSD computed for all resolved PG9 antibody heavy chain C_α atoms using the PyMOL `align` command with the `cycles=0` option.

[2] Numbers in parentheses represent the number of atoms included in the RMSD calculation.

[3] Cryo-EM structure of PG9RSH N(100f)Y.

[4] Cryo-EM structure of DU025.

Table SR6: Flexible residues for designs PGT145 with AMC011 or SIVmac239, related to Results.

PDB ID	**Name**	**Target**	**Mutable res**	**Flexible res**
6NIJ	PGT145:Env (AMC011)	100d, 100l	H100D, H100L	H100A, H100N, A168, A169, C166, C167
8DVD	PGT145:Env (SIVmac239)	100d, 100l	H100D, H100L	H100A, H100N, E168, E169, G166, G167

Table SR5: Predicted K* scores for PGT145 WT and variants using OSPREY vs. corresponding small-panel neutralization data, related to Results. In general, K* scores do not correlate quantitatively with binding affinity[2] except in some cases in which multiple high-resolution structures are used[3]. For these reasons, in this study, the K* scores were used as a filter, to select a small number (10) of candidate PGT145 designs (out of thousands of possible sequences) for experimental testing, with the hypothesis that some would increase breath or potency as predicted.

	IC_{50} (μg/mL)[1]		K* score ($\log_{10}K^*$)[2]		Summary	
Ab Variant	WT	Variant	WT	Variant	$-\Delta\log_{10}IC_{50}$	$\Delta\log_{10}K^*$
F(100d)E	0.080	0.698	5.288	6.911	−0.094	1.623
L(100e)E	0.080	>50	13.763	16.646	< -2.80	2.883
N(100l)D	0.080	0.017	5.288	8.891	0.398	3.603
Y(100m)E	0.080	>50	13.763	15.976	< -2.80	2.213

[1] IC_{50} for BG505.Env.C2 pseudovirus. Geometric mean over several assays reported for wild-type value.

[2] Upper bound on the K* score reported.

Table SR7: Upper and lower bounds on the K* score and partition functions for designs of PGT145 in complex with AMC011 at residues 100d and 100l predicted using OSPREY, related to Results. H_δ and H_ε refer to the δ and ε protonation states of histidine.

		$\log_{10}K^*$		$\log_{10} Z_C$		$\log_{10} Z_{Ab}$		$\log_{10} Z_{Env}$	
100d	100l	UB	LB	UB	LB	UB	LB	UB	LB
W	D	9.845	8.816	46.390	46.149	20.009	19.527	17.324	17.018
Y	D	9.355	8.493	46.396	46.151	20.334	20.023	17.324	17.018
F	D	9.325	8.506	45.941	45.731	19.901	19.598	17.324	17.018
W	Y	9.105	8.096	46.037	45.849	20.429	19.914	17.324	17.018
H_δ	Y	8.857	7.749	44.974	44.759	19.686	19.099	17.324	17.018
D	H_ϵ	8.798	7.841	41.889	41.644	16.479	16.073	17.324	17.018
W	T	8.794	7.720	48.532	48.321	23.277	22.721	17.324	17.018
W	H_ϵ	8.762	7.936	48.123	47.939	22.679	22.342	17.324	17.018
F	H_ϵ	8.737	7.767	47.691	47.491	22.400	21.936	17.324	17.018

Upper and lower bounds on the K* score and partition functions for designs of PGT145 in complex with AMC011 at residues 100d and 100l predicted using OSPREY (continued).

		$\log_{10}$K*		$\log_{10} Z_{\text{C}}$		$\log_{10} Z_{\text{Ab}}$		$\log_{10} Z_{\text{Env}}$	
100d	100l	UB	LB	UB	LB	UB	LB	UB	LB
Y	Y	8.681	7.796	46.114	45.908	20.789	20.415	17.324	17.018
W	Q	8.628	7.512	44.841	44.596	19.760	19.194	17.324	17.018
H_δ	H_ϵ	8.587	7.498	47.009	46.765	21.944	21.404	17.324	17.018
W	S	8.551	7.389	48.141	47.898	23.185	22.572	17.324	17.018
F	Y	8.530	7.681	45.545	45.307	20.302	19.997	17.324	17.018
D	M	8.511	7.389	43.109	42.863	18.151	17.580	17.324	17.018
W	M	8.417	7.427	49.377	49.139	24.388	23.942	17.324	17.018
Y	H_ϵ	8.414	7.547	48.161	47.922	23.051	22.729	17.324	17.018
D	N	8.405	7.380	42.413	42.178	17.475	16.991	17.324	17.018
W	N	8.353	7.418	48.753	48.515	23.774	23.382	17.324	17.018
H_ϵ	H_ϵ	8.253	7.353	46.331	46.092	21.415	21.060	17.324	17.018
F[1]	N[1]	7.941	7.088	48.420	48.174	23.762	23.461	17.324	17.018

[1] Wild-type amino-acid.

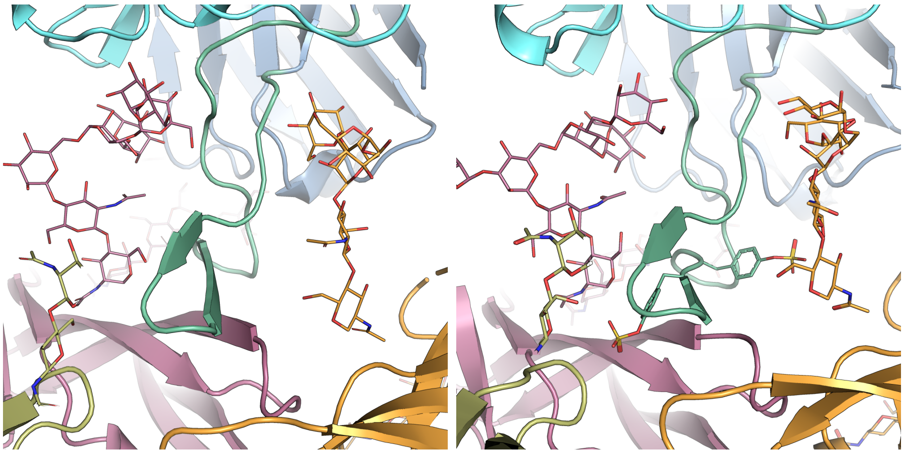
Table SR8: Upper and lower bounds on the K* score and partition functions for designs of PGT145 in complex with SIVmac239 at residues 100d and 100l predicted using OSPREY, related to Results. H_δ and H_ε refer to the δ and ε protonation states of histidine.

		$\log_{10}$K*		$\log_{10} Z_{\text{C}}$		$\log_{10} Z_{\text{Ab}}$		$\log_{10} Z_{\text{Env}}$	
100d	100l	UB	LB	UB	LB	UB	LB	UB	LB
H_δ	E	12.040	10.673	44.537	44.292	21.204	20.585	12.415	11.912
F	E	11.704	10.418	45.280	45.034	22.202	21.664	12.415	11.912
E	M	11.685	10.329	46.220	45.975	23.232	22.624	12.415	11.912
H_ϵ	E	11.504	10.172	45.117	44.872	22.285	21.701	12.415	11.912
E	H_ϵ	11.479	10.210	46.581	46.336	23.712	23.190	12.415	11.912
E	Q	11.418	10.080	45.046	44.801	22.306	21.716	12.415	11.912
E	V	11.311	9.959	44.408	44.163	21.789	21.185	12.415	11.912
E	T	11.250	9.890	44.781	44.536	22.232	21.620	12.415	11.912
E	N	11.243	9.938	45.009	44.764	22.411	21.855	12.415	11.912
E	S	11.153	9.790	44.176	43.931	21.726	21.112	12.415	11.912
E	C	11.099	9.829	44.518	44.273	22.029	21.508	12.415	11.912
H_δ	M	10.970	9.620	47.177	46.932	24.897	24.296	12.415	11.912
F	M	10.946	9.635	48.032	47.786	25.737	25.174	12.415	11.912
F	H_ϵ	10.915	9.600	49.488	49.242	27.228	26.660	12.415	11.912
W	H_ϵ	10.881	9.573	47.387	47.141	25.153	24.593	12.415	11.912
W	N	10.802	9.447	45.429	45.184	23.323	22.716	12.415	11.912
H_δ	H_ϵ	10.773	9.449	48.280	48.035	26.171	25.596	12.415	11.912
Y	R	10.761	9.414	47.990	47.745	25.916	25.318	12.415	11.912
H_δ	Q	10.639	9.286	46.724	46.479	24.778	24.173	12.415	11.912
F	Q	10.539	9.289	47.731	47.486	25.782	25.281	12.415	11.912

Upper and lower bounds on the K* score and partition functions for designs of PGT145 in complex with SIVmac239 at residues 100d and 100l predicted using OSPREY (continued).

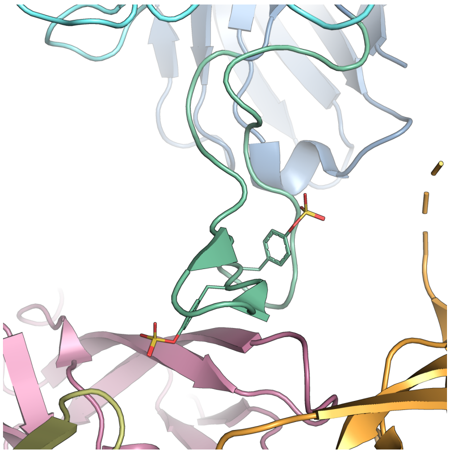
		$\log_{10}$K*		$\log_{10} Z_{\mathrm{C}}$		$\log_{10} Z_{\mathrm{Ab}}$		$\log_{10} Z_{\mathrm{Env}}$	
100d	100l	UB	LB	UB	LB	UB	LB	UB	LB
F[1]	N[1]	10.288	9.235	47.248	47.003	25.353	25.048	12.415	11.912

[1] Wild-type amino-acid.



(A) The CDRH3 loop for PG9RSH N(100f)Y exhibits a twisted conformation.

(B) The CDRH3 loop for DU025 exhibits a twisted conformation.



(C) The CDRH3 loop for PG9 (PDB ID: 5VJ6) exhibits an untwisted conformation.

Figure SR2: The CDHR3 loop conformation differs between PG9RSH variants and PG9, related to Results. Backbone shown in ribbon representation with glycans and sulfated tyrosines shown as lines with Env subunits colored with warm colors, antibody in cool colors, and antibody CDRH3 loops (residues 95-102) shown in green. A pronounced twist in the CDRH3 loop of PG9RSH variants N(100f)Y (A) and DU025 (B) can be seen, contrasted with the untwisted conformation in PG9 (C). As a result, residues 100i - 100n are displaced by up to 3.6 Å.

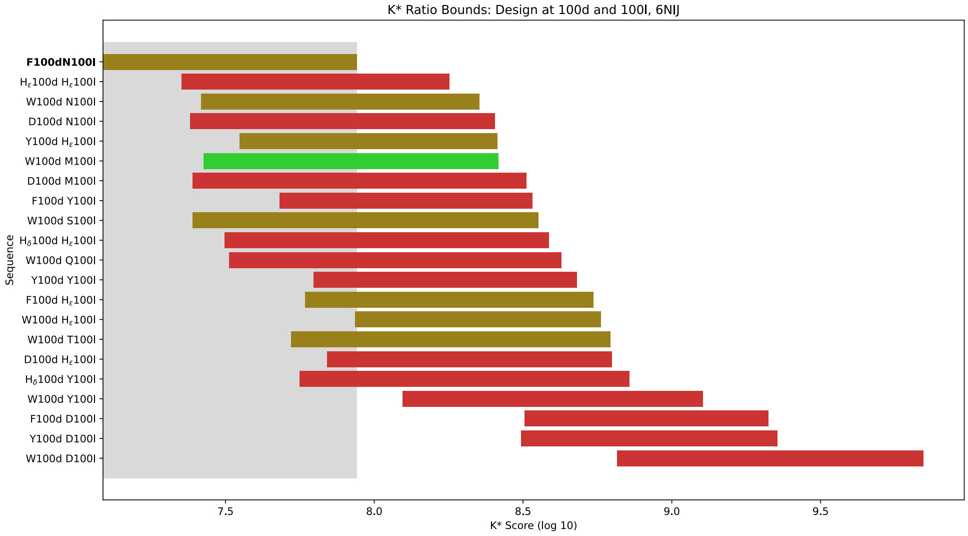


Figure SR3: K* scores for a double-mutation design of PGT145 in complex with AMC011 at residues 100d and 100l, related to Results. Bounds on the K* score for single mutations predicted using OSPREY are shown as horizontal bars. Pictured results are limited to the top 20 design predictions. Designs for which the unbound antibody is predicted to be more stable, approximately equally-stable, or less stable than wild-type are indicated by green, tan, or red coloring, respectively. Relative stability was estimated using the lower-bound on the partition-function value for the unbound antibody state. Wild-type antibody residue labeled in bold. The bounds on the K* score for the wild-type design are indicated by the gray, shaded box.

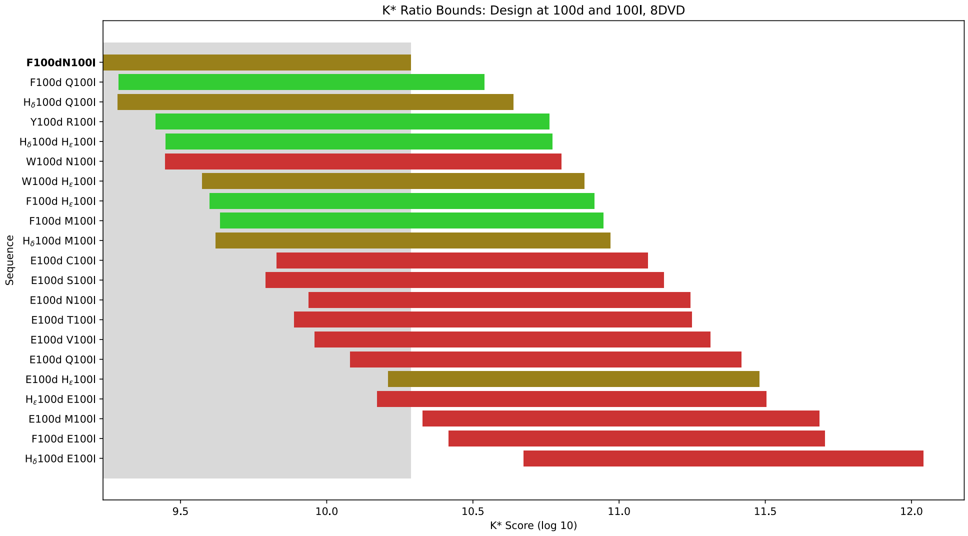
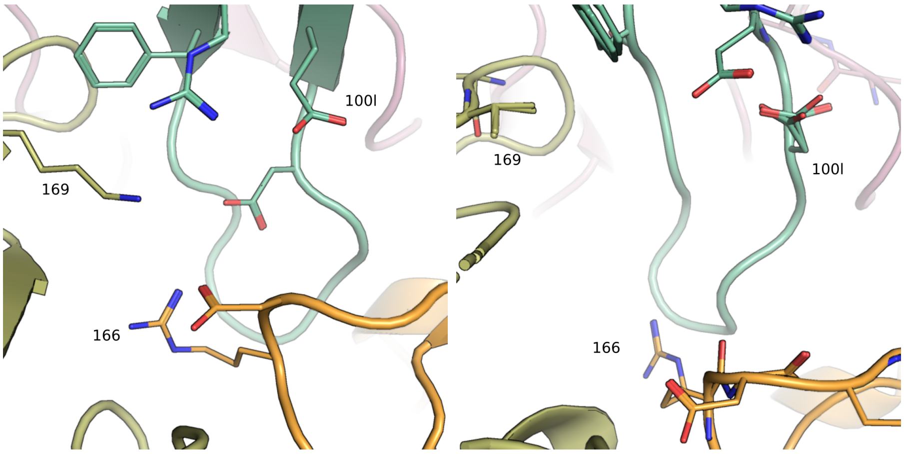
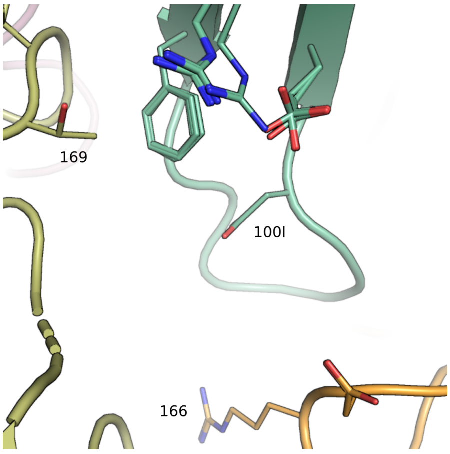


Figure SR4: K* scores for a double-mutation design of PGT145 in complex with SIVmac239 at residues 100d and 100l, related to Results. Bounds on the K* score for single mutations predicted using OSPREY are shown as horizontal bars. Pictured results are limited to the top 20 design predictions. Designs for which the unbound antibody is predicted to be more stable, approximately equally-stable, or less stable than wild-type are indicated by green, tan, or red coloring, respectively. Relative stability was estimated using the lower-bound on the partition-function value for the unbound antibody state. Wild-type antibody residue labeled in bold. The bounds on the K* score for the wild-type design are indicated by the gray, shaded box.

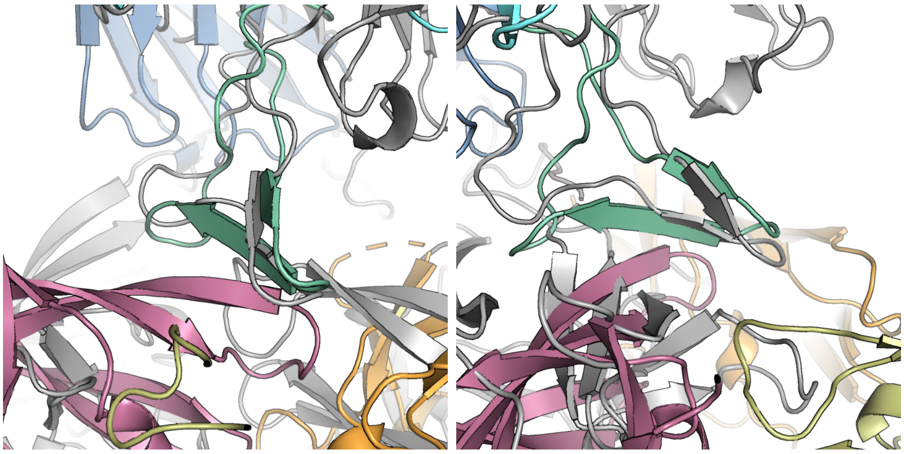


(A) DU303:BG505 Cryo-EM structure. **(B)** DU303:AMC011 model.



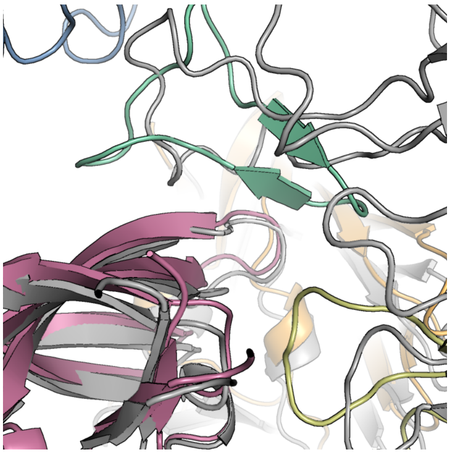
(C) DU303:SIVmac239 model.

Figure SR5: Conservative homology models predict that the N(100l)D mutation will not create intermolecular contacts with AMC011 or SIVmac239, related to Results. Backbone shown in ribbon representation with flexible residues shown as lines with Env subunits colored with warm colors and antibody CDRH3 loops (residues 95-102) shown in green. Models predict that antibody residue D(100l) does not interact with position 169 due to residue substitution, and does not interact with residue 166 due to CDRH3 loop and V2 loop conformational shifts.



(A) AF2 models of DU011 bound to JRFL.JB Env

(B) AF2 models of DU011 bound to ZM233.6 Env



(C) AF2 models of DU025 bound to JRFL.JB Env

Figure SR6: Alphafold Multimer predicted interactions between antibodies and Env are incorrect, related to Results. Backbone shown in ribbon representation. The Cryo-EM structures are shown in color, with Env subunits in warm colors and antibody in cool colors, and antibody CDRH3 loops (residues 95-102) shown in green. Alphafold Multimer models are shown in gray. Interestingly, in some cases Alphafold Multimer does predict that PG9 binds to the Env apex, but incorrectly predicts A) an antiparallel β-strand interaction, B) a "cross strand" interaction, or C) an interaction in which the antibody lies parallel to the Env apex plane. In most cases, including all predicted models with BG505 Env and predicted models of DU025 in complex with TRO.11, physically reasonable predicted structures did not place the antibody at the Env apex (not pictured).

SR1 Extended design results

SR1.1 Validation of design pipeline on published experimental data

To validate our computational protein design methods on the antibody structure input models we first attempted to compare point-mutation designs to published experimental data. We performed four single-residue designs on the PGT145 CDRH3 loop (Residues 100d, 100e, 100l, and 100m) using OSPREY and computed K^* scores for both the wild-type amino-acid and substitution with alanine. We compared the resulting change in K^* score with previously determined IC_{50} data from alanine-scanning of PGT145[1]. The changes in the $\log_{10}\mathrm{K}^*$ scores qualitatively recapitulated trends seen in the experimental data (See Table SR1).

SR1.2 Additional designs for PGT145

We performed double-mutation designs of PGT145 at residues L(100e) and Y(100m) and predicted mutations to improve binding (Figure SR1 and Table SR2). The $\log_{10}\mathrm{K}^*$ score for the wild-type sequence was bounded by the interval $[12.3, 13.8]$ and the $\log_{10} Z_{\mathrm{Ab}}$ was bounded by the interval $[13.8, 14.3]$. Sequences for which the lower-bound on $\log_{10} Z_{\mathrm{Ab}}$ was less than 9.70 were excluded from consideration. Combinations of L(100e)E, L(100e)M, and L(100e)H with Y(100m), Y(100m)M, Y(100m)H, and Y(100m)E were predicted to improve binding without a large reduction in antibody stability.

SR1.3 Additional designs for PG9RSH

Designs of PG9RSH based on a high-resolution, monomeric input model of PG9 binding to the gp120 V1 and V2 loops (PDB ID: 3U2S) returned mutations at various residues predicted to improve binding. A three-residue design of residues R(100b), N(100c), and Y(100e) suggested that positions 100b and 100e are not amenable to mutation, while position 100c exhibits many mutations that are predicted to be slightly favorable, including Met, His, Arg, and Lys. Most mutations are predicted to result in only a slight increase in $\log_{10}\mathrm{K}^*$ score over the wild-type ($\log_{10}\mathrm{K}^* = 13.0$)

SR2 Small-panel neutralization results

We assayed selected designs for neutralization on a small panel of 10 pseudovirus strains. Designs that performed well on these small assays were selected for large-panel (208 pseudoviruses) neutralization assays. Variants were selected for further characterization based on the number of pseudoviruses neutralized with an $\mathrm{IC}_{80} \leq 50\mu\mathrm{g/mL}$, the median IC_{80} value, the number of pseudoviruses neutralized with an $\mathrm{IC}_{50} \leq 50\mu\mathrm{g/mL}$, and the median IC_{50} value (listed in order of importance). Additionally, the sequence and neutralization diversity of the set of variants to be characterized was considered. For example, although several variants (N(100f)Y, DU012-DU017) achieved good neutralization breadth and median potency, only N(100f)Y and DU017 were chosen for further characterization due to their good performance at the IC_{80} cutoff and their incorporation of two different mutations at the 100f position.

Based on the OSPREY designs of PGT145 we selected 10 variants to test in small-panel pseudovirus neutralization assays (See Table S5). These PGT145 variants were assayed for neutralization against a panel of 10 pseudovirus strains from clades A, B, and C (See Data S1). Variant DU303 (containing the N(100l)D mutation) achieved a significant increase in breadth against this small panel, neutralizing 7 of the 10 tested strains with both IC_{50} and $\mathrm{IC}_{80} \leq 50$

μg/mL, while the wild-type controls neutralized only 4 strains (median) with $IC_{50} \leq 50$ μg/mL and only 3 strains (median) with $IC_{80} \leq 50$ μg/mL. The median IC_{50} (0.218 μg/mL) over all neutralized pseudoviruses for DU303 was comparable to that for the wild-type controls, while the median IC_{80} (1.83 μg/mL) was greater than that for the controls.

Based on the OSPREY PG9RSH design results we selected 34 variants to test in small-panel pseudovirus neutralization assays (See Table S6). These variants were assayed for neutralization against a panel of 10 pseudovirus strains from clades A, B, and C (See Data S1). Variant PG9RSH N(100f)Y (containing the N(100f)Y mutation) achieved a clear increase in potency against this small panel, achieving median IC_{50} values of 0.073 and 0.185 μg/mL and median IC_{80} values of 0.821 and 3.87 μg/mL. PG9RSH N(100f)Y also showed a slight increase in breadth as one of the few tested variants to achieve measurable neutralization against the 3873.v1.c24 pseudovirus (clade C). Similarly, variant DU025 (containing the Y(100k)D mutation) achieved an increase in potency against this small panel, with median IC_{50} (0.185 μg/mL) and IC_{80} (2.21 μg/mL) improving relative to the wild-type Ab controls.

Overall, the small-panel neutralization panels indicated that DU303 resulted in increased breadth, PG9RSH N(100f)Y resulted in improved potency with perhaps a slight increase in breadth, and DU025 resulted in improved potency. In summary, the small panels were used as a screening tool to select computationally designed antibodies for further analysis using a large panel. Antibodies with exceptional biological properties, typically neutralization potency or breadth as per the large panel, were subsequently selected for further characterization, including high resolution cryo-EM structures.

SR3 Extended structural analysis of antibody variants

SR3.1 Structural analysis of DU303

The EM structure of PGT145 variant DU303 reveals a binding mode consistent with previous structures of PGT145. We focus our detailed structural analysis around DU303 residue 100l. Applying different masking parameters to the image dataset resulted in two electron density maps with 3.58 Å resolution that differ near residue 100l. The atomic model fit to the second density map shows the antibody interacting with the Env trimer apex through a long, negatively charged CDRH3 (Figure 2C).

D(100l) interacts with residues from two distinct gp120 subunits. Both electron density maps show well-resolved electron density for gp120 residues R166 and K169 (Figure S14). In contrast, density for residue D(100l) and for neighboring F(100d) is more ambiguous. Both maps show significant density between the modeled side-chain locations of D(100l) and F(100d). The first map localizes the density more closely to F(100d), a feature that suggests multiple occupied rotamers for F(100d) (Figure S14A). Density isosurfaces (1.2 and 3.0σ) for residue D(100l) show little density associated with the modeled location of the side-chain carboxyl group, perhaps indicative of a mobile side-chain. The second map, in contrast, localizes the density more closely to residue D(100l) (Figure S14B). This does result in density localized to the modeled location of the side-chain carboxyl of D(100l), but the peak extends further toward the modeled locations of F(100d) and gp120 K169.

The atomic model of DU303 fit to the second density map indicates that D(100l) could form electrostatic interactions with gp120 residues R166 and K169 (Figure 3D). The side-chain nitrogen of K169 lies 5.1 Å from a side-chain carboxyl oxygen of D(100l). Similarly, one of the side-chain nitrogens of R166 lies 4.0 Å from a side-chain carboxyl oxygen of D(100l). The position of these side-chains suggests that the negatively charged D(100l) forms favorable interactions with positively charged residues on gp120 to improve breadth and potency of neutralization.

SR3.2 Comparative analysis with DU303 design ensemble

To determine the extent to which design predictions corresponded with experimental structures we compared 10 members of the low-energy ensemble (LEE) generated during OSPREY design with the new structure of DU303 (Figure 3C,F). We aligned structures by the backbone heavy-atoms of the V2 region of a single gp120 subunit (Chain G, residues 154-177). Members of the LEE aligned to the Cryo-EM model with an RMSD of 0.79 Å calculated using 93 of 96 backbone heavy-atoms.

The CDRH3 loop in the experimental structure model is shifted laterally by 2.1 Å compared to the structures in the LEE, measured at the C_α carbon of residue 100l. Despite this, the distances between the side-chains of D(100l), R166, and K169 are accurately predicted in the LEE. The closest side-chain nitrogen of K169 in the LEE lies 4.2 Å from a side-chain carboxyl oxygen of D(100l). Similarly, one of the side-chain nitrogens of R166 lies 5.0 Å from a side-chain carboxyl oxygen of D(100l). While the LEE predicts that D(100l) interacts more closely with K169 than R166 (a relationship that is inverted in the experimental structure model), these predicted interaction distances differ by at most one angstrom from those in the experimental model. The side-chain orientations of this system are qualitatively similar between the LEE and the experimental model, indicating that OSPREY correctly predicted the structural consequences of the N(100l)D substitution.

SR3.3 Structural analysis of PG9RSH N(100f)Y

The EM structure of PG9RSH N(100f)Y shows the antibody interacting with the Env trimer apex through a hammer-like, negatively charged CDRH3 similar to previous bound structures of PG9RSH (Figure 2A). However, in contrast to a previous electron microscopy structure (PDB ID: 5VJ6)[4], the CDRH3 loop appears in a twisted conformation (Figure SR2). We focus our detailed structural analysis around PG9RSH N(100f)Y residue 100f.

Y(100f) interacts with side-chains and glycans from two distinct gp120 subunits. Electron density maps show well-resolved density for gp120 residues D167, K168, and K169 (Figure S15A). The first two N-acetylglucosamine (GlcNAc) sugars of gp120 glycan N160 are also resolved. Density corresponding to bNAb residues is more ambiguous. Both 1.2 σ and 3.0 σ isosurfaces show density peaks between the modeled side-chain locations of residues Y(100f) and Y(100a), perhaps suggesting the presence of alternate rotamer configurations. These isosurfaces do not contain the modeled locations of terminal hydroxyls for either tyrosine. Examination of low-density peaks (0.5 σ) reveals a small peak in density of the second GlcNAc of glycan N160 that is oriented toward the modeled location of Y(100f), suggesting possible interactions between Y(100f) and the glycan shield (Figure S15B).

The atomic model of PG9RSH N(100f)Y fit to the density map indicates that the primary interaction between Y(100f) and gp120 is a π-cation interaction with residue K168 (Figure 3A). The ammonium nitrogen of K168 lies 4.7 Å from the center of the Y(100f) π system, and the angle between the distance vector and the ring normal vector is approximately 20°, which is representative of typical π-cation geometry[5]. Y(100f) also forms van der Waals interactions with antibody residues P99, Y(100a), and TYS(100h). The side-chain geometry suggests that the aromatic Y(100f) side-chain participates in a π-cation interaction with the positively charged K168 to improve breadth and potency of neutralization.

SR3.4 Comparative analysis with PG9RSH N(100f)Y design ensemble

To determine the extent to which design predictions corresponded with experimental structures we compared 10 members of the low-energy ensemble (LEE) generated during OSPREY design with the new structure of PG9RSH N(100f)Y (Figure 3A, D). We aligned structures by the backbone heavy-atoms of the V2 region of a single gp120 subunit (Chain I, residues 154-177). Members of the LEE aligned to the Cryo-EM model with an RMSD of 0.54 Å calculated using 92 of 96 backbone heavy-atoms.

The CDRH3 loops for the LEE structures and the experimental model near residue 100f are nearly identical in backbone conformation, despite the global difference between the twisted and untwisted conformations. The V2 loop of one gp120 subunit (chain G) exhibits a slight lateral shift relative to the experimental model. The side-chain locations in the LEE are qualitatively similar to those in the experimental model. The ammonium nitrogen lies as close as 5.1 Å from the center of the Y(100f) π system. Interestingly, multiple rotamers of Y(100f) and Y(100a) appear in the LEE, resulting in a conformation in which these side-chains have rotated and stacked. While for the majority of the LEE Y(100f) presents the face of the π ring toward K168, for a few members K168 forms van der Waals interactions with the edge of the aromatic ring. Notably, the design ensemble indicates D167 may interact with Y(100f), but these predicted interactions are not supported by the experimental model. However, the overall correspondence between LEE and experimental model suggests that OSPREY correctly predicted the structural consequences of the N(100f)Y substitution.

SR3.5 Structural analysis of DU025

The EM structure of PG9RSH variant DU025 shows the antibody interacting with the Env trimer apex through a hammer-like, negatively charged CDRH3 similar to previous bound structures of PG9RSH (Figure 2A). However, in contrast to a previous electron microscopy structure (PDB ID: 5VJ6)[4], the CDRH3 loop appears in a twisted conformation (Figure SR2).

We focus our detailed structural analysis around DU025 residue 100k. The electron density around the side chain of D(100k) is well-resolved at 1.2σ and maintains a small but noticeable peak at 3.0σ (Figure S16A). Similarly, clear density is observed at both 1.2 and 3.0σ for the side-chains of Env residues Q170, K305, and Y173, along with the first two core GlcNAc monomers of glycan N156. Nearly the entire core of glycan N156 is resolved at 1.2σ. Interestingly, three unassigned density peaks arise in the groove between the V2 and V3 loops at both 1.2 and 3.0σ, which could indicate the presence of solvent at this interface (Figure S16B). Furthermore, a bridge of density at 1.0σ arises between the modeled locations of Env residues Q170 and R308, hinting at long-range or solvent-mediated interactions between the two. Multiple strong density peaks around the modeled side-chain location of R308 provide evidence for multiple side-chain conformations at this residue.

These data suggest that residue D(100k) may form long-range or solvent-mediated interactions with residues Q170 and K305, which lie at distances of 4.1, and 6.8 Å, respectively. Additionally, D(100k) may interact with residue Y173 and glycan N156. The multiple unsatisfied densities at both low and high σ suggest that solvent may play a critical role at this interface.

SR3.6 Comparative analysis with DU025 design ensemble

To determine the extent to which design predictions corresponded with experimental structures we compared 10 members of the low-energy ensemble (LEE) generated during OSPREY design with the new structure of DU025 (Figure 3B, E). We aligned structures by the backbone heavy-atoms of the V2 region of a single gp120 subunit (Chain I, residues 154-177). Members of the LEE aligned to the Cryo-EM model with an RMSD of 0.50 Å calculated using 90 of 96 backbone heavy-atoms.

Members of the LEE indicate that improvement in binding is predicted to result from favorable electrostatic interactions between D(100k) and Q170 or R308. The lowest-energy member of the design ensemble places one carboxyl oxygen of D(100k) 3.2 Å from the side chain nitrogen of Q170 and 5.6 Å from R308. K305, on the other hand, lies distal (9 Å) from D(100k).

The design ensemble (based on PDB: 5VJ6) displays the CDRH3 loop in the untwisted conformation, whereas the cryo-EM structure of DU025 exhibits a twisted CDRH3 loop. Although this does not cause large changes for most PG9RSH residues that contact Env directly, residues 100i-100n are displaced by up to 3.6 Å measured at the C_α. Residue D(100k) in particular is displaced by 3.5 Å, resulting in a significant change in its environment. As a result, although the design ensemble correctly predicts that D(100k) creates long-range electrostatic interactions, and correctly predicts a favorable interaction with Q170, the predicted interaction between D(100k) and R308 is not supported by the cryo-EM structure. Instead, shifts in the backbone create interactions between D(100k) and K305. However, the overall quality and type of interactions formed by D(100k) in the design ensemble is in fact consistent with the experimental structure.

SR3.7 Inconsistent CDRH3 loop conformation in design inputs of PG9

Interestingly, we noticed differences between the CDRH3 conformations of PG9 of our design input model (based on PDB ID: 5VJ6) and the experimentally determined structures of both PG9RSH N(100f)Y and DU025. Cryo-EM structures placed the PG9RSH CDRH3 loop in a twisted conformation, similar to the conformations displayed in previous structures (PDB IDs: 3MME, 3U2S, 3U4E, 4YAQ), but the model represented by 5VJ6 displays an “untwisted” conformation (Figure SR2). The consensus between all but one structure suggests that the twisted conformation is in fact the correct CDRH3 conformation. The aberrant structure, 5VJ6, was generated by fitting crystal structures of PG9 bound to a V1V2 scaffold (PDB ID: 3U4E) and unliganded gp120 core (PDB ID: 3DNN) to density obtained via electron microscopy[6], but surprisingly the loop conformations in 5VJ6 and 3U4E are different. This suggests that fitting or processing of 5VJ6 inappropriately modified the CDRH3 loop, resulting in a hallucinated conformation.

SR4 Comparison to putative germline antibodies

The mutation defining PGT145 DU303, N(100l)D, does not appear in the putative germline antibody ([7] via IMGT/V-QUEST[8]) nor the clonally-related PGT141-PGT144 sequences[9]. As such, this mutation appears to represent a higher bar for elicitation by vaccines, albeit by only one mutation. It is important to note that, due to the long CDRH3 insertion, the CDRH3 sequence of the putative germline antibody was taken from the mature PGT145 antibody when the D gene did not provide sufficient coverage[7].

Similarly, the mutations PG9RSH N(100f)Y and PG9RSH DU025 (Y(100k)D) do not appear in the putative germline antibodies for PG9 ([7] via IMGT/V-QUEST[8]) nor the somatically-related PG16[10]. Nor do either of these mutations appear in the mature PG16 antibody. These observations suggest that, like DU303, these mutations represent a higher bar for elicitation. However, position 100k does appear contain a mutation away from germline in both PG9 and PG16, suggesting that, at least for DU025, the bar for vaccine elicitation is no higher than for the wild-type PG9 antibody, ignoring the RSH chimeric context.

SR5 Investigation of additional Env genotypes

We are interested in additional questions and extensions to this work, but unfortunately many of these are beyond the scope of the current manuscript. Can this method be applied to *any* antigen genotype to successfully improve breadth? We suspect that the answer is no. We have in fact been careful to avoid claims that improved breadth can result from design against *any* antigen in this manuscript. How, then, can we differentiate between antigens that are more "canonical" (i.e., antigens for which design for potency can result in breadth), versus those that are less? In the discussion section of this manuscript we provide several hypotheses for how one could begin to address this question, but we leave the evaluation of these hypotheses for future work. Designs on additional genotypes could provide information to help answer these questions, although structural information for these antibodies complexed with other Env genotypes is sparse. The results of such designs would not themselves be definitive, and design results would need to be characterized experimentally.

Nonetheless, to satisfy our curiosity we additionally peformed OSPREY designs on PGT145 in complex with the AMC011 (PDB ID: 6NIJ) and SIVmac239 (PDB ID: 8DVD) Env trimers, focusing on CDRH3 residues 100d and 100l. The primary differences between these trimers and the BG505 trimer at this site lie at Env residue 169, which is a lysine, valine, or threonine for the BG505, AMC011, and SIVmac239 trimers, respectively. We used the unminimized input PDB structures (after adding hydrogens using MolProbity) as the input structure, with flexible and mutable residues described in Table SR6, and approximated the top 20 sequences by K^* score (and wild-type sequence) for each model.

The results of these designs can be found in the Supplementary Information (Table SR7 and SR8, Figure SR3 and SR4). Overall, we do see sequences with high K^* scores within the top 20 sequences of both designs that appear similar to those in the original BG505 design, especially when focusing on the mutations at position 100d. Differences can be seen with respect to postion 100l, which is expected given the sequence variation at residue 169. Similarly to the BG505 designs, both the 6NIJ and 8DVD designs assign high K^* scores to sequences with either the N(100l)D or N(100l)E mutations, and, like the BG505 design, the AMC011 design assigns a high K^* score to the F(100d) D(100l) sequence. However, the low antibody partition function Z_{Ab} and low complex partition function for the PGT145-F(100d)D(100l):AMC011 sequence indicates that this is not a viable mutation for binding to the AMC011 Env trimer.

Most substitutions of N(100l) with a negatively charged residue result in a decrease in the unbound antibody partition function Z_{Ab} to some extent, indicating that the additional negative charge may decrease the antibody CDRH3 stability. However, examining the complex and antibody partition functions for these sequences reveals a critical difference: The complex partition function for the BG505:PGT145-D100l complex remains approximately equal to that for the BG505:PGT145, but the complex partition function for the AMC011:PGT145-D100l complex is significantly lower than that for the AMC011:PGT145 complex. Thus, our models predict that the the negatively charged D(100l) is compensated for in the BG505 complex state by interaction with K169, but not in the AMC011 complex state, where no productive interactions are formed with V169. The high K^* score of F(100d)-D(100l) in the AMC011 design is a case of a bad complex partition function divided by a worse unbound partition function – a phenomenon that we select against using the unbound partition function stability cutoffs. As a result, we would not have chosen to test the N(100l)D mutation from this design, but would have identified position 100l as a site of interest.

Overall, these additional designs suggest that OSPREY is correctly modeling the environment at this site. Similar mutations were found by OSPREY ranked by K^* score alone, but there were significant differences in results after filtering by stability. It is impossible to assess the effects of many of these mutations (e.g., Y(100d)W, N(100l)M) without experimental characterization, but these results provide a direction for possible future work.

SR6 Structure modeling of additional Env genotypes

To gain further insight into the mechanisms of improved breadth for these variant antibodies we constructed homology models of each antibody in complex with Env trimers from other strains of HIV or SIV. We first leveraged existing structures of PGT145 in complex with the AMC011 (PDB ID: 6NIJ) and SIVmac239 (PDB ID: 8DVD) trimers to construct conservative homology models. Unfortunately, to our knowledge there are still no structures of PG9 in complex with Env trimers from strains other than BG505. The primary difference between AMC011 or SIVmac239 trimers and the BG505 trimer near residue 100l is the identity of Env residue 169, which is a lysine, valine, or threonine in BG505, AMC011, and SIVmac239, respectively. We used the unminimized input PDB structures (after adding hydrogens using MolProbity) as the input structure, with flexible and mutable residues described in Table SR6. We used OSPREY to mutate PGT145 residue N(100l) to aspartate in both structures and generated ensembles of predicted low energy conformations in the same manner as described in the methods section of the manuscript. We then compared the resulting ensembles to the Cryo-EM structure of DU303 to model the effect of Env sequence and structural variation on DU303.

Broad structural features match well between the models of DU303 bound to AMC011 or SIVmac239 and the DU303:BG505 experimental structure, but small differences in loop conformations are evident (See Figure SR5) We aligned structures by the 23 C_α atoms of the CDRH3 loops, which aligned to the cryo-EM structure of DU303:BG505 with an RMSD of 1.62 Å or 0.86 Å for the models of DU303:AMC011 or DU303:SIVmac239, respectively. Slight differences are evident in the CDRH3 loop of DU303:AMC011 relative to DU303:BG505, which significantly change the orientation of the C_α-C_β vector of residue 100l relative to the DU303:BG505 structure. The CDRH3 loop of DU303:SIVmac239 appears quite similar to that of DU303:BG505. The V2 loops in both models are roughly consistent with those in the experimental structure, but shifts of up to 3 Å or 7 Å relative to the BG505 V2 conformations can be seen. These differences could reflect real conformational differences, conformational heterogeneity, model fitting error, or experimental uncertainty.

The intermolecular interactions of residue D(100l) in both models are significantly reduced relative to the experimental structure. Not only is V169 too short to interact with the relatively distant D(100l) in the DU303:AMC011 model, but the reorientation of residue 100l due to CDRH3 conformation differences seems to preclude positive interactions with V169. Additionally, this reorientation increases the distance between residue 100l and R166, and, as a result, the model predicts no significant interactions between these residues. T169 is also too short to interact with D(100l) in the DU303:SIVmac239 model. R166 appears to interact primarily with the sulfated tyrosine at residue 100f, possibly due to shifts in the V2 loop position relative to the DU303 experimental structure. Overall, these models suggest that DU303 would not gain neutralization potency for the AMC011 HIV strain and for the SIVmac239 SIV strain due to lack of interaction with the important K169 residue. Similar observations were also previously gleaned from the analysis of large-panel neutralization data (See Figure S10).

Due to the lack of experimental structure information for PG9 binding to Env trimers other than BG505, to model the interactions of P9RSH N(100f)Y and DU025 we also pursued less conservative modeling approaches. We used AlphaFold Multimer[11,12] to predict structures of PG9RSH N(100f)Y in complex with Envs from strains ZM233.6, and JRFL.JB, and predicted structures of DU025 in complex with JRFL.JB and TRO.11. Additionally, we predicted structures of all antibodies in complex with BG505.W6M.C2 as a positive control. We examined 25 predicted structures for each complex. In general, although the antibody and Env structures individually showed good homology to existing structures, the antibody binding epitope was poorly predicted. Only a few predicted structures placed antibodies near the Env apex (Figure SR6), and no structures successfully recapitulated the parallel β-strand interaction between PG9RSH and the Env V2 loop. Interestingly, no predicted structures for the BG505 control placed these antibodies near the apex of the Env trimer.

Overall, our modeling yielded good results when based on conservative modification of existing experimental structures, but was less successful when such structures were unavailable. Due to the scarcity of experimental structure data, the difficulty of modeling protein loops, and the difficulty of accurately predicting binding interfaces, further modeling of the interactions between these variant antibodies and diverse Env trimers is deferred to future work.

References

[1] Jeong Hyun Lee et al. "A Broadly Neutralizing Antibody Targets the Dynamic HIV Envelope Trimer Apex via a Long, Rigidified, and Anionic β-Hairpin Structure". In: *Immunity* 46.4 (Apr. 2017), pp. 690–702. ISSN: 10974180. DOI: 10.1016/j.immuni.2017.03.017. URL: https://pubmed.ncbi.nlm.nih.gov/28423342/.

[2] Mark A. Hallen et al. "OSPREY 3.0: Open-Source Protein Redesign for You, with Powerful New Features". In: *J. Comput. Chem.* 9999 (2018), pp. 1–14. DOI: 10.1002/jcc.25522. URL: https://www.biorxiv.org/content/early/2018/04/23/306324.

[3] Anna Ulrika Lowegard. "Novel Algorithms and Tools for Computational Protein Design with Applications to Drug Resistance Prediction, Antibody Design, Peptide Inhibitor Design, and Protein Stability Prediction". PhD thesis. Duke University, 2022.

[4] Haoqing Wang et al. "Asymmetric recognition of HIV-1 Envelope trimer by V1V2 loop-targeting antibodies". In: *eLife* 6 (2017).

[5] Hervé Minoux and Christophe Chipot. "Cation-π interactions in proteins: Can simple models provide an accurate description?" In: *Journal of the American Chemical Society* 121.44 (Nov. 1999), pp. 10366–10372. ISSN: 00027863. DOI: 10.1021/ja990914p. URL: http://www.sander.embl-heidelberg.de/.

[6] Jean Philippe Julien et al. "Asymmetric recognition of the HIV-1 trimer by broadly neutralizing antibody PG9". In: *Proceedings of the National Academy of Sciences of the United States of America* 110.11 (Mar. 2013), pp. 4351–4356. ISSN: 00278424. DOI: 10.1073/pnas.1217537110. URL: https://www.pnas.org/content/110/11/4351%20https://www.pnas.org/content/110/11/4351.abstract.

[7] Kwinten Sliepen et al. "Binding of inferred germline precursors of broadly neutralizing HIV-1 antibodies to native-like envelope trimers". In: *Virology* 486 (Dec. 2015), pp. 116–120. DOI: 10.1016/j.virol.2015.08.002. URL: https://doi.org/10.1016%2Fj.virol.2015.08.002.

[8] Véronique Giudicelli, Xavier Brochet, and Marie-Paule Lefranc. "IMGT/V-QUEST: IMGT Standardized Analysis of the Immunoglobulin (IG) and T Cell Receptor (TR) Nucleotide Sequences". In: *Cold Spring Harb Protoc* 2011.6 (June 2011), pdb.prot5633. DOI: 10.1101/pdb.prot5633. URL: https://doi.org/10.1101%2Fpdb.prot5633.

[9] Laura M. Walker et al. "Broad neutralization coverage of HIV by multiple highly potent antibodies". In: *Nature* 477.7365 (Sept. 2011), pp. 466–470. ISSN: 00280836. DOI: 10.1038/nature10373. URL: https://www.nature.com/articles/nature10373.

[10] Marie Pancera et al. “Crystal Structure of PG16 and Chimeric Dissection with Somatically Related PG9: Structure-Function Analysis of Two Quaternary-Specific Antibodies That Effectively Neutralize HIV-1”. In: *Journal of Virology* 84.16 (Aug. 2010), pp. 8098–8110. ISSN: 0022-538X. DOI: 10.1128/jvi.00966-10. URL: http://jvi.asm.org/.

[11] John Jumper et al. “Highly accurate protein structure prediction with AlphaFold”. In: *Nature* 596.7873 (Aug. 2021), pp. 583–589. ISSN: 14764687. DOI: 10.1038/s41586-021-03819-2. URL: https://doi.org/10.1038/s41586-021-03819-2.

[12] Richard Evans et al. “Protein complex prediction with AlphaFold-Multimer”. In: *bioRxiv* (Mar. 2022), p. 2021.10.04.463034. DOI: 10.1101/2021.10.04.463034. URL: https://doi.org/10.1101/2021.10.04.463034.