

**Structural survey of HIV-1-neutralizing antibodies targeting Env trimer delineates epitope categories and suggests vaccine templates**

Gwo-Yu Chuang<sup>a,1,2</sup>, Jing Zhou<sup>a,1</sup>, Reda Rawi<sup>a</sup>, Chen-Hsiang Shen<sup>a</sup>, Zizhang Sheng<sup>b</sup>,  
Anthony P. West, Jr.<sup>c</sup>, Baoshan Zhang<sup>a</sup>, Tongqing Zhou<sup>a</sup>, Robert T. Bailer<sup>a</sup>, Nicole A. Doria-Rose<sup>a</sup>,  
Mark K. Louder<sup>a</sup>, Krisha McKee<sup>a</sup>, John R. Mascola<sup>a</sup>, Pamela J. Bjorkman<sup>c</sup>, Lawrence Shapiro<sup>a,b</sup>,  
and Peter D. Kwong<sup>a,b,2</sup>

<sup>a</sup> Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA.

<sup>b</sup> Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032, USA.

<sup>c</sup> Division of Biology, Howard Hughes Medical Institute, California Institute of Technology, Pasadena, California 91125, USA.

<sup>1</sup> G.-Y.C. and J.Z. contributed equally to this work.

<sup>2</sup> To whom correspondence should be addressed: Email: gwo-yu.chuang@nih.gov, pdkwong@nih.gov

Running title: Trimer recognition by HIV neutralizing antibodies

Key words: antibody recognition, B cell ontogeny, broadly neutralizing antibody, cryo-electron microscopy, envelope glycoprotein trimer, glycan shielding, sequence variation, vaccine design, x-ray crystallography

**HIV-1 broadly neutralizing antibodies are desired for their therapeutic potential and as templates for vaccine design. Such antibodies target the HIV-1-envelope (Env) trimer, which is shielded from immune recognition by extraordinary glycosylation and sequence variability. Recognition by broadly neutralizing antibodies thus provides insight into how antibody can bypass these immune-evasion mechanisms. Remarkably, antibodies neutralizing >25% of HIV-1 strains have now been identified that recognize all major exposed surfaces of the prefusion-closed Env trimer. Here we analyzed all 206 broadly neutralizing antibody-HIV-1 Env complexes in the PDB with resolution suitable to define their interaction chemistries. These segregated into 20 antibody classes based on ontogeny and recognition, and into 6 epitope categories (*V1V2, glycan-V3, CD4-binding site, silent face center, fusion peptide, and subunit interface*) based on recognized Env residues. We measured antibody neutralization on a 208-isolate panel and analyzed features of paratope and B cell ontogeny. The number of protruding loops, CDR H3 length, and level of somatic hypermutation for broadly HIV-1 neutralizing antibodies were significantly higher than for a comparison set of non-HIV-1 antibodies. For epitope, the number of independent sequence segments was higher ( $P < 0.0001$ ), as well as the glycan component surface area ( $P = 0.0005$ ). Based on B cell ontogeny, paratope, and breadth, the CD4-binding site antibody IOMA appeared to be a promising candidate for lineage-based vaccine design. In terms of epitope-based vaccine design, antibody VRC34.01 had few epitope segments, low epitope-glycan content, and high epitope-conformational variability, which may explain why VRC34.01-based design is yielding promising vaccine results.**

## **Significance**

Over the last decade, structures of broadly neutralizing antibodies have been determined, encompassing all major exposed surfaces on the prefusion-closed HIV-1-envelope (Env) trimer. To provide insight into how the immune system generated these antibodies along with molecular detail of their recognition, we surveyed known Env-antibody structures. We found that (i) broadly neutralizing antibodies recognized sequence variability and glycan at frequencies similar to their average occurrence on the Env-trimer surface, (ii) almost two-thirds of the paratopes utilized protruding loops, which were otherwise less common; and (iii) the broadest HIV-1-neutralizing antibodies had high levels of SHM, and the most potent unusual recombination, but antibodies with more reproducible B cell ontogenies still neutralized broadly, and these may be better templates for re-elicitation.

Antibodies against HIV-1 that neutralize a significant fraction of the diverse primary isolates that typify HIV-1 transmission are highly sought. These antibodies target the HIV-1 envelope (Env) trimer, which is comprised of gp120 and gp41 subunits and shielded from immune recognition by extraordinary glycosylation (1), sequence variability (2), and conformational masking (3). A decade ago, only four such broadly neutralizing antibodies had been identified. However, in 2009, the identification by B cell culture of antibody PG9 (along with a somatic variant PG16) (4) and in 2010 the identification by probe-based sorting of antibody VRC01 (along with somatic variants VRC02 and VRC03) (5) initiated an outbreak of discovery, with dozens of broadly neutralizing antibodies identified by culture or probe-based sorting of B cells from HIV-1 infected donors (reviewed in (6)).

Importantly, structures for many of these antibodies in complex with HIV-1 Env have been determined, which link antibody function (breadth and potency of neutralization) with molecular features of antibody recognition (paratope) and chemical details of Env-recognized site (epitope). For example, structures of PG9 with scaffolded V1V2 regions of Env (7) and also with Env trimer (8) have revealed a protruding anionic tyrosine-sulfated loop penetrating the glycan shield to interact with a conserved cationic site on V2; similarly structures of VRC01 with core gp120 (9) or with a fully glycosylated Env trimer (10) have revealed how SHM-based optimization of interactions – especially those involving glycan – allows for broad recognition of the CD4-binding site. Most recently, the structure of antibody VRC-PG05 in complex with a glycosylated Env core (11) has shown that even the highly glycosylated center of the Env “silent face” could be recognized by broadly neutralizing antibodies.

The structure of VRC-PG05 in complex with Env completes recognition by structurally characterized antibody of all major exposed surfaces on the prefusion-closed HIV-Env trimer.

Here we take advantage of this completion to perform a structure-based survey of all antibody-HIV-1 Env structures in the Protein Data Bank (PDB). We limited analyses to structures of human antibodies (with neutralization breadth > 25%) in complex with epitopes on the prefusion-closed Env trimer with resolutions sufficient to define side chain orientation and interactions. The focus on the prefusion-closed conformation of Env alleviated issues of conformational masking and enabled comparisons on a single structural entity. We further classified structures by antibody class (B cell ontogeny and structural mode of recognition) to unique classes and used a bottom-up classification based on residues of each epitope to define epitope categories. We measured antibody neutralization on a diverse cross-clade panel of 208 HIV-1 isolates and correlated structural features of epitope and paratope with functional characteristics of neutralization breadth and potency. We also analyzed paratopes for ease of lineage re-elicitation and epitopes for ease of mimicry. Overall, our structural survey of HIV-1 antibodies targeting the prefusion-closed Env trimer delineates categories, identifies underlying relationships between structural properties of paratope and epitope and functional properties of neutralization, and suggests favorable candidates for vaccine design.

## **Results**

### **Broadly neutralizing antibodies that recognize the prefusion-closed Env trimer segregate into 20 classes and 6 categories**

The PDB contains over 200 coordinate sets for human antibodies in complex with various portions of HIV-1 Env, ranging from fully glycosylated Env trimers to core gp120, scaffolded domains, representative *N*-linked glycans, and peptide fragments (**Dataset S1**). We chose to limit our analysis to structures with resolutions sufficient to define chemistry of

interaction: for X-ray crystal structures, we used a resolution limit of 3.9 Å, and for cryo-electron microscopy structures, we used a resolution limit of 4.5 Å. Also, we chose to limit our analysis to antibodies that neutralize at least 25% of HIV-1 at 50 µg/ml, as measured in a cross-clade panel of at least 100 isolates. Finally, to alleviate issues of conformation, we limited our analysis to only those antibodies that recognize the prefusion-closed conformation of Env as this allowed us to focus on a single structural entity: the Env trimer in its prefusion-closed state.

We next used similarity in B cell ontogeny and structural mode of recognition to classify these PDBs into antibody classes (12) (**Table 1**). While some antibody classes have only been found in a single donor to date (e.g., CH103, VRC38), other classes have been found in more than one donor (e.g. VRC01, 8ANC131). The largest class involved the VRC01 class (13), with 53 PDBs describing 33 antibodies. Antibodies recognizing glycan-V3, derived primarily from classes that have only a single known clonal lineage, although PGT121 and BG18 lineages do share the same D3-3 gene and mode of recognition (14, 15). For antibodies recognizing the spike apex, prior analyses placed antibody lineages PG9, PGT145 and CH01 in the same class (7). While structural information showed parallel strand recognition for PG9 and CH03 indicating similar classes (16), structures of PGT145 with Env trimer (17, 18) indicated loop insertion into a trimeric hole at the spike apex, and hence a distinct mode of recognition and class for PGT145. We used the name of the first identified antibody in each antibody class as the name of the class.

We selected a structural representative antibody for each class based on the most informative structure (for example, choosing an antibody-Env trimer complex over the same antibody with a deglycosylated core gp120) (**Fig. 1A**) Based on the 20 representative antibody-antigen complex structures, we defined epitope residues on the prefusion-closed Env trimer by buried surface area calculations (see methods). For 2G12, the most complete antibody-Env

complexes defined interactions with only a few *N*-linked glycans (19); therefore, we used published biochemical studies to define the 2G12 epitope residues (20). Next, we carried out hierarchical clustering of epitope residues, which clustered the 20 antibody epitopes into 6 categories (**Fig. 1A,B**). Three of the categories had been previously described: V1V2, glycan-V3 and CD4-binding site, and each of these categories encompassed at least three different antibody classes. Three additional categories were identified, including the “silent face center” (with antibody VRC-PG05), “fusion peptide” (with antibodies PGT151 and VRC34), and “subunit interface” (with antibodies 35O22 and 8ANC195). Beyond shared epitope residues, each of the categories did appear to have specific characteristics. For example, smFRET measurements define three prefusion states (states 1-3) of the Env trimer and found broadly neutralizing antibodies to preferentially recognize state 1 (21); only two antibodies, VRC34.01 (22) and PGT151 (personal communications, Walther Mothes), have thus far been found to shift the smFRET population to state 2, and both were assigned to the same epitope category. Thus, our structural analysis defined six categories of broadly neutralizing antibodies, and antibodies in each category generally shared similar characteristics.

### **Neutralization characteristics for 20 antibody classes**

As the most important functional property of the 20 antibody classes is their ability to neutralize diverse HIV-1 isolates, we felt it important to measure accurately this feature in a way that allowed comparison between different antibodies. We thus measured both neutralization breadth and potency for the representative member of each class on a diverse cross-clade panel of 208 HIV-1 isolates (23) (**Fig. 2A,B, Data Set S2**). For functional representative, we chose the antibody-class member with highest neutralization breadth. For the PGT121 class, we chose

antibody 10-1074 (14); for the VRC01 class, we chose antibody N6 (24); and for the 8ANC131 class we chose antibody CH235.12 (25). Notably, the newly measured median neutralization IC<sub>50</sub> generally differed by less than 2 fold from prior reported values (5, 7, 11, 22, 24-40), with the breadth for all 20 classes averaging 56.7%, with an average geometric mean IC<sub>50</sub> of 0.27 µg/ml.

### **B cell ontogenies, paratopes, and epitopes of antibodies recognizing the prefusion-closed Env trimer**

B cell ontogenies involve the recombination events that lead to creation of the ancestor B cell for each lineage, as well as SHM that lead to mature antibody clones. To estimate elicitation likelihoods related to recombination and SHM, we extracted CDR H3 length and V-gene SHM (**Fig. 2C**). We observed that the average CDR H3 length and heavy chain SHM of the 20 HIV-1 broadly neutralizing antibodies were higher than those of next generation sequencing (NGS) reads from three healthy donors ( $P < 0.0001$ ) (25).

We also extracted critical features of the paratope (**Fig. 2D**). We selected a measure related to recognition mode (number of protruding antibody loops) and a second measure related to recognition extent (the surface area of the interacting paratope). Thirteen of the 20 antibodies utilized protruding loops in recognition, although this feature was less common in non-HIV-1 antibodies (as defined on a set of 81 non-HIV-1 antibody-antigen structures, see **Datasets S3** and **S4**). The paratope surface areas were almost twice as high as observed with non-HIV-1 antibodies ( $P < 0.0001$ ), and this may relate to glycan interactions, as antibodies with deglycosylated cores generally showed lower interactive surface areas.



For epitope, we measured several features including sequence conservation, extent of glycan contribution, number of independent sequence elements in each epitope, and the structural variability of the epitope (**Fig. 3A-D**). For sequence conservation, epitopes of broadly neutralizing antibodies (75% conservation) were similar to the average conservation of the closed trimer (69%). For glycan, we observed the average contribution to epitope surface area (45%) was similar to the surface area contribution of MAN5 glycans to the overall Env trimer surface area (53%). Also, we observed that the average epitope glycan composition ( $P=0.0005$ ), total epitope surface area ( $P<0.0001$ ), and number of epitope segments ( $P<0.0001$ ) of the 20 HIV-1 broadly neutralizing antibodies were higher than those of non-HIV-1 antibodies (**Datasets S3 and S4**; for glycan composition specifically, the comparison was done with a non-HIV-1 glycoprotein set with 16 antibody antigen complexes, see Methods and **Dataset 5**). It may be that features common in HIV-1 neutralizing antibodies, but generally rare, provide insight into recognition constraints: in this case suggesting that HIV broadly neutralizing antibodies were selected specifically to overcome sequence variability and glycosylation.

### **Underlying relationships between neutralization, paratope and epitope**

To provide insight into other features specific to these HIV-1 neutralizing antibodies, we correlated neutralization with epitope and paratope features to reveal underlying relationships (**Fig. 4A**). Although a number of features correlated with p-values of less than 0.05, when corrected for multiple comparison, few correlations were significant. Significant correlations included the expected positive correlation between surface area of paratope and surface area of epitope ( $P < 0.0001$ ,  $r = 0.9845$ ) (**Fig. 4B**) and a negative correlation between protein and glycan surface areas ( $P = 0.0002$ ,  $r = -0.8413$ ) (**Fig. 4C**). This negative correlation suggests antibodies to

have an overall surface area that can comprise either protein or glycan; thus, a glycan-focused antibody generally has lower protein-interactive surface than a protein-focused antibody, and vice versa.

Other correlations did not achieve statistical significance, but were potentially revealing. For example, we observed antibody neutralization breadth to correlate positively with degree of heavy chain SHM ( $P = 0.0510$ ,  $r = 0.4538$ ) and buried protein-epitope surface area ( $P = 0.0079$ ,  $r = 0.5892$ ) and negatively with degree of epitope glycan ( $P = 0.0105$ ,  $r = -0.6579$ ) (**Fig. 4D**), indicating that while antibody recognition can include glycan, this nevertheless reduces breadth. Meanwhile sequence conservation correlated with size of the epitope ( $P = 0.0254$ ,  $r = 0.6961$ )/paratope ( $P = 0.0211$ ,  $r = 0.7110$ ) (**Fig. 4E**), suggesting that one way to overcome variation is by reducing the area of recognized surface. We also noticed that the broadest HIV-1-neutralizing antibodies had high levels of SHM and the most potent had unusual CDR H3s (**Fig. 2**). Overall, these observations suggested both the frequency and the extent that neutralizing antibodies targeting the prefusion-closed Env trimer utilize protruding loops, unusual recombination, and extensive SHM to overcome immune evasion mechanism of extraordinary glycosylation and high sequence variability.

### **Re-elicitation: Antibody-lineage based vaccine design**

We also analyzed antibodies to identify which B cell ontogenies might be most easily re-elicited by antibody-lineage based vaccine design. Such design is premised on the re-elicitation of antibodies with similar ontogenies, through priming of ancestor B cells and induction of their maturation (16, 41, 42). Two factors affect re-elicitation: the likelihood that an appropriate recombination event produces an appropriate unmutated common ancestor (UCA) and the

likelihood that this UCA will mature through processes of SHM to achieve similar functional properties of neutralization. The former is affected by prevalence of appropriate V-genes and by the contribution to the paratope of CDR H3 features such as D-gene and N-nucleotide addition, which we calculated for each of the 20 classes of neutralizing antibodies (**Fig. 5**). We also calculated features of SHM, such as the percent of the V gene-contributed paratope surface altered by SHM or the prevalence of rare mutations (43). Although some of the broadest and most potent antibodies had features that lowered the probability of their re-elicitation, there were antibody classes with less rare SHM and CDR H3 characteristics such as IOMA, as well as antibodies like PG9 with extensive D-gene contributions to antigen recognition that might be re-elicited with greater ease.

## **Discussion**

The completion by antibody VRC-PG05 of structural characterization of the recognition by broadly neutralizing antibodies of all major exposed surfaces on the prefusion-closed Env trimer enables both comprehensive and aggregate-level analyses: Relative to paratope, we could address which characteristics allow for broad recognition of the glycosylated and sequence-variable Env trimer, and relative to the Env trimer surface, we could address which features were recognized and which were avoided. For paratope, our results indicated protruding loops and extensive SHM to dominate recognition (an observation previously reported by Klein and colleagues (44) utilizing a less complete set of HIV-1-directed antibodies), although antibodies with more common features could be found. For epitope, our results indicate both glycan and sequence variation to be recognized by antibody, which may relate to the extent that these features dominate the surface of the prefusion-closed Env trimer.

While our focus on the prefusion-closed trimer allowed us to relate properties of recognized epitopes to average properties of the Env trimer in a specific conformation, this focus led to the omission from our analysis of an important class of broadly neutralizing antibodies, those that target the membrane-proximal external region (MPER) (**Fig. S1** and **Table S2**). MPER antibodies generally show low levels of recognition of the native Env trimer (45) and have thus been proposed to recognize Env trimer in a different conformation (46). In addition to differences in conformation of Env recognized, MPER antibodies also target regions of Env with significantly lower glycan and sequence variability. Thus, the immune-evasion mechanisms that shield the MPER appear to differ from those shielding the prefusion-closed Env trimer, and the antibody mechanisms that allow for recognition of the MPER likely differ as well from those that allow recognition of the closed Env trimer.

In terms of antibody templates most favorable for vaccine design, our analysis suggested the IOMA lineage as being most favorable for re-elicitation by antibody lineage-based approaches due to the lack of unusual features in its B cell ontogeny. Furthermore, in comparison with VRC01 class antibodies, IOMA may overcome the N276-glycan barrier more readily: CDR L1 in IOMA (and in most other antibodies from VL2-23 and related germline genes) has a short  $\alpha$ -helix that can accommodate the N276 glycan with fewer mutations than required by VRC01 class antibodies (47). In terms of features most favorable to epitope-based vaccine design, we note that antibody VRC34.01 had few epitope segments, low epitope-glycan content, and high epitope-conformational variability (48). These features may suggest that vaccine design based on the epitope of antibody VRC34.01 may be especially promising. Indeed, we recently achieved the elicitation of antibodies in mice, guinea pigs and rhesus macaques that neutralized diverse HIV-1 Env isolates through immunogen design based on the epitope of VRC34.01 (49). Thus, in

addition to revealing features of paratope and epitope that allow for immune recognition of the prefusion closed spike, the structural survey also provides insight into which antibody templates are most suitable for vaccine design.

## **Materials and Methods**

### **Structural Dataset and Selection of Structural Representative for Each of the 20 Unique**

**HIV-1 Antibody Classes.** All the antibody-antigen complex structures were collected from Protein Data Bank (PDB) (50) using the key words “HIV-1 antibody” on 28 February 2018.

Structures that contain antibody alone were excluded. Non-human and synthetic antibodies were also excluded. The final comprehensive list of current deposited HIV-1 human antibody-antigen complex structures is provided in **Dataset S1**. Only the antibodies that target envelope trimer neutralize Tier 2 HIV-1 isolates were considered. To focus on broadly neutralizing HIV-1 antibodies, a neutralizing breadth cutoff of >25% was also used. Based on ontogenies in separate donors, we defined 20 unique classes that recognize the prefusion-closed Env trimer (**Fig. 1A**).

For each antibody class, we selected the most informative structure based on the following criteria: 1) Resolution cutoff for X-ray crystal structures < 3.9 Å and < 4.5 Å for cryo-EM structures, allowing for a margin of 0.5 Å in resolution if a high resolution antibody-epitope complex were also available. 2) The complex structure of the trimer with highest resolution was selected, followed by glycosylated monomer, by deglycosylated monomer, or by scaffolded domain with highest resolution. The final PDB IDs we selected included 5V8L (18) (for PGT145), 3U2S (7) (for PG9), 5VGJ (34) (for VRC38.01), 5CEZ (51) (for 3H+109L and 35O22), 5ACO (52) (for PGT128), 4JM2 (53) (for PGT135), 1OP5 (19) (for 2G12), 4YE4 (for HJ16), 5T3Z (47) (for IOMA), 5F9O (25) (for CH235.09), 5FYJ (10) (for VRC01), 4YDJ (32) (for VRC13.01), 5VN8 (54) (for b12), 4JAN (28) (for CH103), (32) 4YDK (for VRC16.01),

6BF4 (11) (for VRC-PG05), 5FUU (55) (for PGT151), 5I8H (22) (for VRC34.01), and 5CJX (56) (for 8ANC195).

**Epitope and Paratope Buried Surface Area Calculations.** The buried surface area between antibody and antigen was calculated using program NACCESS (57). The epitope residues and paratope residues for each antibody were defined as residues with non-zero buried surface area. For 2G12, the epitope residues were defined as glycans N295, N332, and N392, based on (20).

**Structural Definition of HIV-1 Env Broadly Neutralizing Antibody Categories.** The structural representative of the 20 unique antibodies were clustered using hierarchical clustering, with the distance between each of the antibody is defined as  $1 - \frac{n_o}{N_s}$ , where  $n_o$  is the number of overlapping epitope residues, and  $N_s$  is the number of epitope residues for the antibody with the least number of epitope residues of the two.

**Display of Epitope Residues on Prefusion-Closed HIV-1 Env.** The BG505 SOSIP trimer structure (PDB:5FYL) was used as the template with each glycan (with the addition of glycans 241 and 332) modeled as MAN5. Amino acid and glycan epitope residues were colored based on the structurally defined antibody category. Epitope residues with more than 5 Å C-alpha deviation from PDB:5FYL were excluded. For residue positions shared in epitopes from multiple antibody categories, the color of the antibody category with the antibody that has the largest epitope surface area was used.

**Antibody Expression and Purification.** The 20 representative antibody heavy and light chain expression constructs were synthesized (Gene Universal Inc., Newark, DE) and cloned into pVRC8400 expressing vector. For antibody protein expression, 1.5 ml of Turbo293 transfection reagent (Speed BioSystems) were mixed into 25 ml Opti-MEM medium (Life Technology) and incubated for 5 min at room temperature. 500 µg of plasmid DNAs (250 heavy chain and 250 µg

of light chain) were mixed into 25 ml of Opti-MEM medium in another tube. Then the diluted Turbo293 were added into Opti-MEM medium containing plasmid DNAs. Transfection reagent and DNA mixture was incubated for 15 min at room temperature, and added to 400 ml of Expi293 cells (Life Technology) at 2.5 million cells/ml. The transfected cells were cultured in shaker incubator at 120 rpm, 37 °C, 9% CO<sub>2</sub> overnight. On the next day of transfection, 40 ml of AbBooster medium (ABI scientific) were added to each flask of transfected cells and the flasks were transferred to shaker incubators at 120 rpm, 33 °C, 9% CO<sub>2</sub> for additional 5 days. At 6 days post transfection, antibodies in clarified supernatants were purified over 3 ml Protein A (GE Health Science) resin in columns. Antibody was eluted with IgG elution buffer (Pierce), immediately neutralized with one tenth volume of 1M Tris-HCL pH 8.0. The antibodies were then buffer exchanged in PBS by dialysis, adjusted concentration to 1.0 mg/ml and filtered (0.22 µm) for neutralization assays (58) (45).

**Virus Neutralization.** Single-round-of-replication Env pseudoviruses were prepared, titers were determined, and the pseudoviruses were used to infect TZM-bl target cells as described previously in an optimized and qualified automated 384-well format (59). Briefly, antibodies were serially diluted, a constant amount of pseudovirus added, and plates incubated for 60 minutes; followed by addition of TZMbl cells which express luciferase upon viral infection. The plates were incubated for 48 hours and then lysed, and luciferase activity measured. Percent neutralization was determined by the equation:  $(\text{virus only}) - (\text{virus} + \text{antibody}) / (\text{virus only})$  multiplied by 100. Data are expressed as the antibody concentration required to achieve 50% neutralization (IC<sub>50</sub>) and calculated using a dose-response curve fit with a 5-parameter nonlinear function. We used a previously described panel (5, 34, 60) of 208 geographically and genetically diverse Env pseudoviruses representing the major subtypes and circulating recombinant forms.

The IC<sub>50</sub> values reported here are from the complete set of 208 viruses run at the VRC. In some cases, multiple runs were averaged. The values used here differed slightly, generally within 2 fold of previous median IC<sub>50</sub>, from previously reported values in references (5, 7, 11, 22, 24-40). The differences have two sources: in earlier publications, the panels contained between 178 and 198 of the 208 viruses; and because the neutralization IC<sub>50</sub> values are known to vary up to 3-fold between repeat assays, as documented in (59).

### **CDR H3 Length, Germline Gene Assignment, N-Insertion, and SHM Calculations.**

Antibody CDR H3 length, germline gene assignment, N-insertion, and SHM Calculation was performed using the IMGT/HighV-QUEST webservice (<http://www.imgt.org/HighV-QUEST/>).

**Number of Protruding Loops.** CDR loops (i.e. CDR L1, CDR L2, CDR L3, CDR H1, CDR H2, CDR H3) were inspected manually, and two criteria were used in defining whether a CDR loop would be considered as protruding: 1) the loop extends at least 5Å beyond a plane defined by the average extent of the six antibody CDRs, and 2) the loop contacts with HIV epitope residues.

**Epitope Conservation Calculations.** The conservation of each HIV-1 Env residue was calculated using the entropy function of the R package bio3d (H.norm column), with the addition of glycan as residue type. The calculation was based on the filtered web alignment of the year 2016 from the Los Alamos HIV sequence database (<http://www.hiv.lanl.gov/>). The epitope conservation was defined as the average conservation for each residue, weighted by epitope surface area. The average Env surface conservation was defined as the average conservation for each Env residue, weighted by accessible surface area as determined by NACCESS (57).

**Average Glycan Surface Calculations.** The glycan surface area of the HIV-1 Env prefusion-closed trimer was calculated applying a two-step procedure. First, we approximated the



accessible surface area (ASA), using NACCESS program (57) with the default radius of 1.4 Å, for all protein as well as glycan residues based on a molecular dynamics simulation with a 500-nanosecond timescale (10). Second, we determined the percentage of the glycan surface area, by dividing the ASA of glycan residues by the sum of the ASAs of protein and glycan residues.

**Number of Epitope Segment Calculations.** The number of epitope segment is defined as number of continuous stretches on epitope residues, allowing skipping of one to two residues, which in total have an epitope surface area of greater than the specified threshold ( $25\text{\AA}^2$  or  $100\text{\AA}^2$ ).

**Epitope Structural Variation Calculations.** The epitope structural variation for each antibody class was determined by the averaged root-mean-square-deviation of recognized epitope amino acid residues computed pairwise among category members and the unliganded structure (PDB ID: 4ZMJ) (33) using PyMOL software. Antibody 2G12 conformational variability was not calculated due to lack of amino acid epitope residues.

**Correlation Analyses.** Pearson correlation for every pair of antibody properties listed in **Fig. 2** and **3** were calculated. b12 class was not included in the correlation calculation as b12 was derived from phage display. For correlations involving total epitope surface, paratope surface, and number of epitope segment calculations, PG9, VRC38.01, PGT135, H16, CH235, VRC13.01, CH103, VRC16.01, and VRC-PG05 were excluded as their representative structures were determined with partial Env domain. For correlations involving epitope glycan composition, HJ16, CH235, VRC13.01, CH103, and VRC16.01 were excluded as their representative structures were determined with deglycosylated gp120 monomers.

**Non-HIV-1 Antibody Sequence Dataset and V-gene Frequency Estimation.** Heavy and light chain NGS samples (43) from 454 pyrosequencing downloaded from Short Read Archive and

processed to calculate average SHM (**Table S1A – S1C**) and HV germline frequency (**Table S1D**). The next-generation sequencing (NGS) data from healthy donor performed with 5' primer-adapters should avoid bias of germline gene usage. All heavy chain reads shorter than 350 nucleotides were removed, and all lambda and kappa chain reads longer than 300 nucleotides were kept. Germline genes were assigned to all filtered reads using IgBLAST (61). After assigned all reads, an in-house python script applied to process IgBLAST output, and non-IG sequences were removed. The variation between alleles were ignored, and we treated those alleles as identical VH gene. The VH gene frequency was calculated by the number of reads of VH germline divided by the total good IG sequences.

**SHM Rarity.** Rarity score of a somatic hypermutation was calculated as 1-frequency of the SHM observed in gene-specific substitution profiles (43). SHMs with rarity score lower than 0.5% were counted as rare SHMs.

**Non-HIV Antibody-Antigen Structural Dataset.** A non-redundant non-HIV-1 Env antibody/antigen structural dataset was obtained from SAbDab (62), using the “Non-redundant search” function (Antibody identity = 90%, Antigen identity = 99% , for protein antigens, Resolution = 3.9 Å). The list was further filtered down by keeping only the human antibodies, and removing all HIV antibodies, resulting a total of 81 antibody/antigen complexes (**Dataset S2**).

**Non-HIV Glycoprotein Structural Dataset.** To understand the difference of epitope glycan composition between HIV-1 antibodies and non-HIV-1 glycoprotein, a non-HIV-1 glycoprotein structural dataset was assembled with 16 antibodies complexed with one of the following glycoproteins: Ebola virus glycoprotein, Herpes simplex viruses -2 gD, Human Metapneumovirus fusion protein, Hepatitis C virus envelope glycoprotein E2 core, dengue virus

type 2 envelope glycoprotein, Human cytomegalovirus glycoprotein B, Respiratory syncytial F, Human cytomegalovirus pentamer, and influenza hemagglutinin (**Dataset S5**). To avoid the dataset over-represented by influenza hemagglutinin, only four structures (the number of structures for the next abundant antigen) were selected. Resolution cutoff for X-ray crystal structures  $< 3.9 \text{ \AA}$  and  $< 4.5 \text{ \AA}$  for cryo-EM structures.

## Acknowledgements

We thank J. Stuckey for assistance with figures, and members of the Structural Biology Section and Structural Bioinformatics Core, Vaccine Research Center, for discussions or comments on the manuscript. We thank H. Mouquet and M. Nussenzweig for the nucleotide sequences of antibody 10-1074. We also thank J. Baalwa, D. Ellenberger, F. Gao, B. Hahn, K. Hong, J. Kim, F. McCutchan, D. Montefiori, L. Morris, J. Overbaugh, E. Sanders-Buell, G. Shaw, R. Swanstrom, M. Thomson, S. Tovanabutra, C. Williamson, and L. Zhang for contributing the HIV-1 envelope plasmids used in our neutralization panel. Support for this work was provided by the Intramural Research Program of the Vaccine Research Center, National Institute of Allergy and Infectious Diseases, and by IAVI Neutralizing Antibody Consortium (NAC).

**Author contributions:** G.-Y.C., L.S. and P.D.K. designed research; G.-Y. C., J.Z., R.R., C.-H. S., Z. S., A. P. W., B.Z., T.Z., R.T.B., N.A.D. and .M.L. performed research; G.-Y.C., L.S., and P.D.K. wrote the paper, with all authors providing comments or revisions.

## References

1. Wei X, *et al.* (2003) Antibody neutralization and escape by HIV-1. *Nature* 422(6929):307-312.

2. Starcich BR, *et al.* (1986) Identification and characterization of conserved and variable regions in the envelope gene of HTLV-III/LAV, the retrovirus of AIDS. *Cell* 45(5):637-648.
3. Kwong PD, *et al.* (2002) HIV-1 evades antibody-mediated neutralization through conformational masking of receptor-binding sites. *Nature* 420(6916):678-682.
4. Walker LM, *et al.* (2009) Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. *Science* 326(5950):285-289.
5. Wu X, *et al.* (2010) Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science* 329(5993):856-861.
6. Burton DR & Mascola JR (2015) Antibody responses to envelope glycoproteins in HIV-1 infection. *Nat Immunol* 16(6):571-576.
7. McLellan JS, *et al.* (2011) Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. *Nature* 480(7377):336-343.
8. Julien JP, *et al.* (2013) Asymmetric recognition of the HIV-1 trimer by broadly neutralizing antibody PG9. *Proc Natl Acad Sci U S A*.
9. Zhou T, *et al.* (2010) Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. *Science* 329:811-817.
10. Stewart-Jones GB, *et al.* (2016) Trimeric HIV-1-Env Structures Define Glycan Shields from Clades A, B, and G. *Cell* 165(4):813-826.
11. Zhou T, *et al.* (2018) A Neutralizing Antibody Recognizing Primarily N-Linked Glycan Targets the Silent Face of the HIV Envelope. *Immunity* 48(3):500-513 e506.
12. Kwong PD & Mascola JR (2012) Human antibodies that neutralize HIV-1: identification, structures, and B cell ontogenies. *Immunity* 37(3):412-425.
13. Zhou T, *et al.* (2013) Multidonor analysis reveals structural elements, genetic determinants, and maturation pathway for HIV-1 neutralization by VRC01-class antibodies. *Immunity* 39(2):245-258.
14. Mouquet H, *et al.* (2012) Complex-type N-glycan recognition by potent broadly neutralizing HIV antibodies. *Proc Natl Acad Sci U S A* 109(47):E3268-3277.
15. Barnes CO, *et al.* (2018) Structural characterization of a highly-potent V3-glycan broadly neutralizing antibody bound to natively-glycosylated HIV-1 envelope. *Nat Commun* 9(1):1251.
16. Gorman J, *et al.* (2016) Structures of HIV-1 Env V1V2 with broadly neutralizing antibodies reveal commonalities that enable vaccine design. *Nat Struct Mol Biol* 23(1):81-90.
17. Liu Q, *et al.* (2017) Quaternary contact in the initial interaction of CD4 with the HIV-1 envelope trimer. *Nat Struct Mol Biol* 24(4):370-378.
18. Lee JH, *et al.* (2017) A Broadly Neutralizing Antibody Targets the Dynamic HIV Envelope Trimer Apex via a Long, Rigidified, and Anionic beta-Hairpin Structure. *Immunity* 46(4):690-702.
19. Calarese DA, *et al.* (2003) Antibody domain exchange is an immunological solution to carbohydrate cluster recognition. *Science* 300(5628):2065-2071.
20. Sanders RW, *et al.* (2002) The mannose-dependent epitope for neutralizing antibody 2G12 on human immunodeficiency virus type 1 glycoprotein gp120. *J Virol* 76(14):7293-7305.
21. Munro JB, *et al.* (2014) Conformational dynamics of single HIV-1 envelope trimers on the surface of native virions. *Science* 346(6210):759-763.

22. Kong R, *et al.* (2016) Fusion peptide of HIV-1 as a site of vulnerability to neutralizing antibody. *Science* 352(6287):828-833.
23. Seaman MS, *et al.* (2010) Tiered categorization of a diverse panel of HIV-1 Env pseudoviruses for assessment of neutralizing antibodies. *J Virol* 84(3):1439-1452.
24. Huang J, *et al.* (2016) Identification of a CD4-Binding-Site Antibody to HIV that Evolved Near-Pan Neutralization Breadth. *Immunity* 45(5):1108-1121.
25. Bonsignori M, *et al.* (2016) Maturation Pathway from Germline to Broad HIV-1 Neutralizer of a CD4-Mimic Antibody. *Cell* 165(2):449-463.
26. Doria-Rose NA, *et al.* (2012) HIV-1 neutralization coverage is improved by combining monoclonal antibodies that target independent epitopes. *J Virol* 86(6):3393-3397.
27. Huang J, *et al.* (2012) Broad and potent neutralization of HIV-1 by a gp41-specific human antibody. *Nature* 491(7424):406-412.
28. Liao HX, *et al.* (2013) Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. *Nature* 496(7446):469-476.
29. Pancera M, *et al.* (2013) Structural basis for diverse N-glycan recognition by HIV-1-neutralizing V1-V2-directed antibody PG16. *Nat Struct Mol Biol* 20(7):804-813.
30. Chuang GY, *et al.* (2013) Residue-level prediction of HIV-1 antibody epitopes based on neutralization of diverse viral strains. *J Virol* 87(18):10047-10058.
31. Rudicell RS, *et al.* (2014) Enhanced potency of a broadly neutralizing HIV-1 antibody in vitro improves protection against lentiviral infection in vivo. *J Virol* 88(21):12669-12682.
32. Zhou T, *et al.* (2015) Structural Repertoire of HIV-1-Neutralizing Antibodies Targeting the CD4 Supersite in 14 Donors. *Cell* 161(6):1280-1292.
33. Kwon YD, *et al.* (2015) Crystal structure, conformational fixation and entry-related interactions of mature ligand-free HIV-1 Env. *Nat Struct Mol Biol* 22(7):522-531.
34. Cale EM, *et al.* (2017) Virus-like Particles Identify an HIV V1V2 Apex-Binding Neutralizing Antibody that Lacks a Protruding Loop. *Immunity* 46(5):777-791 e710.
35. Bonsignori M, *et al.* (2017) Staged induction of HIV-1 glycan-dependent broadly neutralizing antibodies. *Sci Transl Med* 9(381).
36. Alam SM, *et al.* (2017) Mimicry of an HIV broadly neutralizing antibody epitope with a synthetic glycopeptide. *Sci Transl Med* 9(381).
37. Julg B, *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(406).
38. Xu L, *et al.* (2017) Trisppecific broadly neutralizing HIV antibodies mediate potent SHIV protection in macaques. *Science* 358(6359):85-90.
39. Pancera M, *et al.* (2017) Crystal structures of trimeric HIV envelope with entry inhibitors BMS-378806 and BMS-626529. *Nat Chem Biol* 13(10):1115-1122.
40. Williams WB, *et al.* (2017) Initiation of HIV neutralizing B cell lineages with sequential envelope immunizations. *Nat Commun* 8(1):1732.
41. Jardine J, *et al.* (2013) Rational HIV immunogen design to target specific germline B cell receptors. *Science* 340(6133):711-716.
42. Crowe JE, Jr. (2016) Teaching a Clone to Walk, One Step at a Time. *Cell* 166(6):1360-1361.
43. Sheng Z, *et al.* (2017) Gene-Specific Substitution Profiles Describe the Types and Frequencies of Amino Acid Changes during Antibody Somatic Hypermutation. *Front Immunol* 8:537.

44. Klein F, *et al.* (2013) Somatic mutations of the immunoglobulin framework are generally required for broad and potent HIV-1 neutralization. *Cell* 153(1):126-138.
45. Chakrabarti BK, *et al.* (2011) Direct antibody access to the HIV-1 membrane-proximal external region positively correlates with neutralization sensitivity. *Journal of virology* 85(16):8217-8226.
46. Ruprecht CR, *et al.* (2011) MPER-specific antibodies induce gp120 shedding and irreversibly neutralize HIV-1. *J Exp Med* 208(3):439-454.
47. Gristick HB, *et al.* (2016) Natively glycosylated HIV-1 Env structure reveals new mode for antibody recognition of the CD4-binding site. *Nat Struct Mol Biol* 23(10):906-915.
48. Tainer JA, *et al.* (1984) The reactivity of anti-peptide antibodies is a function of the atomic mobility of sites in a protein. *Nature* 312(5990):127-134.
49. Xu K, *et al.* (2018) Epitope-based vaccine design yields fusion peptide-directed antibodies that neutralize diverse strains of HIV-1. *bioRxiv* <https://doi.org/10.1101/306282>.
50. Berman HM, *et al.* (2000) The Protein Data Bank. *Nucleic Acids Res* 28(1):235-242.
51. Garces F, *et al.* (2015) Affinity Maturation of a Potent Family of HIV Antibodies Is Primarily Focused on Accommodating or Avoiding Glycans. *Immunity* 43(6):1053-1063.
52. Lee JH, de Val N, Lyumkis D, & Ward AB (2015) Model Building and Refinement of a Natively Glycosylated HIV-1 Env Protein by High-Resolution Cryoelectron Microscopy. *Structure* 23(10):1943-1951.
53. Kong L, *et al.* (2013) Supersite of immune vulnerability on the glycosylated face of HIV-1 envelope glycoprotein gp120. *Nat Struct Mol Biol* 20(7):796-803.
54. Ozorowski G, *et al.* (2017) Open and closed structures reveal allostery and pliability in the HIV-1 envelope spike. *Nature* 547(7663):360-363.
55. Lee JH, Ozorowski G, & Ward AB (2016) Cryo-EM structure of a native, fully glycosylated, cleaved HIV-1 envelope trimer. *Science* 351(6277):1043-1048.
56. Scharf L, *et al.* (2015) Broadly Neutralizing Antibody 8ANC195 Recognizes Closed and Open States of HIV-1 Env. *Cell* 162(6):1379-1390.
57. Hubbard SJ & Thornton JM (1993) 'NACCESS', Computer Program, Department of Biochemistry and Molecular Biology, University College London.
58. Kwon Y, *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(7):1798-1809.
59. Sarzotti-Kelsoe M, *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.
60. Georgiev IS, *et al.* (2013) Delineating antibody recognition in polyclonal sera from patterns of HIV-1 isolate neutralization. *Science* 340(6133):751-756.
61. Ye J, Ma N, Madden TL, & Ostell JM (2013) IgBLAST: an immunoglobulin variable domain sequence analysis tool. *Nucleic Acids Res* 41(Web Server issue):W34-40.
62. Dunbar J, *et al.* (2014) SAbDab: the structural antibody database. *Nucleic Acids Res* 42(Database issue):D1140-1146.

## Figure legends

### **Fig. 1. 20 classes of broadly neutralizing antibodies recognize the prefusion-closed Env trimer and segregate into 6 categories based on the Env residues with which they interact.**

(A) All HIV-1 Env-antibody complexes structures in the PDB (Table S1) were assigned to classes (leftmost column, listed by the name of first reported antibody of the class) based on similarities in B cell ontogeny and mode of recognition, with a representative structure and PDB for the class (2nd and 3rd columns from left), which were chosen based on resolution and degree to which Env in the structure resembled prefusion-closed trimer; “\*” indicate structures determined in gp120-core context; “#” indicates structures determined with V1/V2 scaffold; “@” indicates glycan N295, N332, and N392 used for 2G12 epitope (see methods). (B) Prefusion-closed Env trimer with molecular surface colored by categories defined in (A). Epitope residues shared by antibodies in separate classes were colored according to surface area and requirements for antibody binding. For example, glycan N276 has interactions with antibody 8ANC195 of the subunit interface category and with multiple antibodies of the CD4-binding site category; however, because glycan N276 is required for 8ANC195, and generally only accommodated by antibodies that target the CD4-binding site, we colored glycan N276 to be part of the subunit interface category. Left image is shown with viral membrane at bottom; right image is rotated 90° to look down on the trimer apex.

### **Fig. 2. Neutralization, B cell ontogeny, and paratope characteristics for 20 classes of broadly neutralizing antibodies.**

(A) The broadest neutralizing antibody from each class was chosen as representative. (B) Neutralization properties. Bar graphs show breadth and potency of each representative antibody on the 208-isolate panel, with dotted line showing averages. (C) Features of B cell ontogeny. CDR H3 lengths are defined based on Kabat nomenclature. @ indicates that antibody b12 is derived from a phage library. IOMA is the only antibody with CDHR H3 length under 20, SHM <12%, and breadth approaching 50%, and thus may be a promising candidate for lineage-based vaccine design. By contrast, VRC01-class antibodies have high SHM and PG9/PGT145 have unusually long CDR H3, complicating their lineage-based elicitation. (D) Paratope properties. “\*” indicate structures determined in gp120-core context; “#” indicates structures determined with V1/V2 scaffold. Structures denoted with “\*” or “#” were included in calculations of the category averages for protruding loop numbers, but not for paratope surface area. P-values were determined using two-tailed Mann-Whitney test.

### **Fig. 3. Epitope features for 20 classes of broadly neutralizing antibodies.**

(A) Categories, classes, and representative structures. “\*” indicate structures determined in gp120-core context; “#” indicates structures determined with V1/V2 scaffold; @ indicates that the epitope structural variation of antibody 2G12 was not calculated due to the lack of protein epitope. Structures denoted with “\*” or “#” were not included in calculations of the category averages for number of epitope segments and epitope total surface. Structures denoted with “\*” were not included in calculations of the category average for epitope glycan composition. (B) Epitope properties. Bar graphs show conservation and glycan-interactive surface for each representative antibody, with dotted line showing averages. “\*” indicate structures determined in gp120-core context, where glycan surface areas may not reflect trimer interactions. (C) Number of epitope segments, shown for segments with  $>25 \text{ \AA}^2$  or  $>100 \text{ \AA}^2$  surface area. VRC34 has low number epitope segments and low glycan context (black arrow), and this may explain why fusion peptide immunization are yielding promising results from epitope-based vaccine design. (D) Observed variation in epitope structure, as calculated from RMSD of recognized amino acid residues, computed pairwise among class members of each category and unliganded trimer. (E) Prefusion-closed Env trimer colored by epitope and by various epitope features. P-values were determined using two-tailed Mann-Whitney test.

#### **Fig. 4. Correlations of neutralization and features of epitope and paratope reveal underlying relationships.**

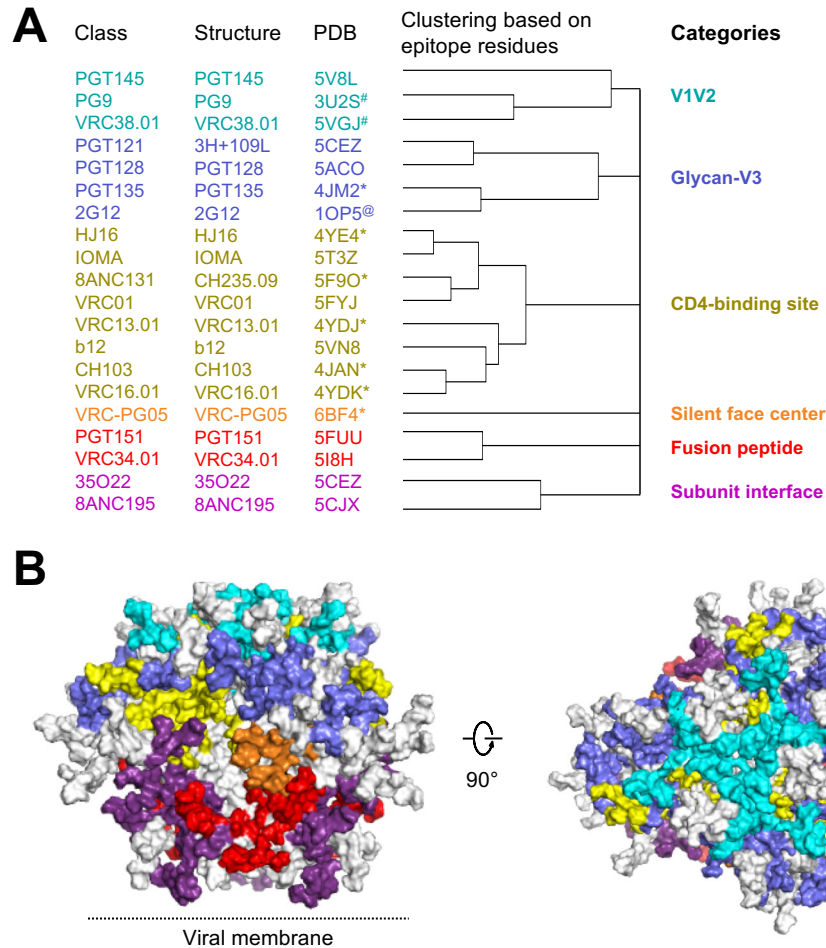
Correlation matrices are shown for properties of neutralizing antibodies that recognize the prefusion closed Env trimer. (A) Pearson correlations ( $r$ ) are provided in the upper right; P-values (not adjusted for multiple comparisons) are provided in the lower left. Correlations with P-value of less than 0.06 were highlighted with dark red font, and correlations with P-value of less than 0.05 after Bonferroni correction were highlighted with both dark red font and orange fill. (B) Correlation between paratope surface area and epitope surface area. (C) Correlation between Epitope protein surface area and glycan component of surface. (D) Correlation between neutralizing breadth and heavy chain SHM. (E) Correlation between epitope conservation and epitope total surface area.

#### **Fig. 5. Factors affecting re-elicitation.**

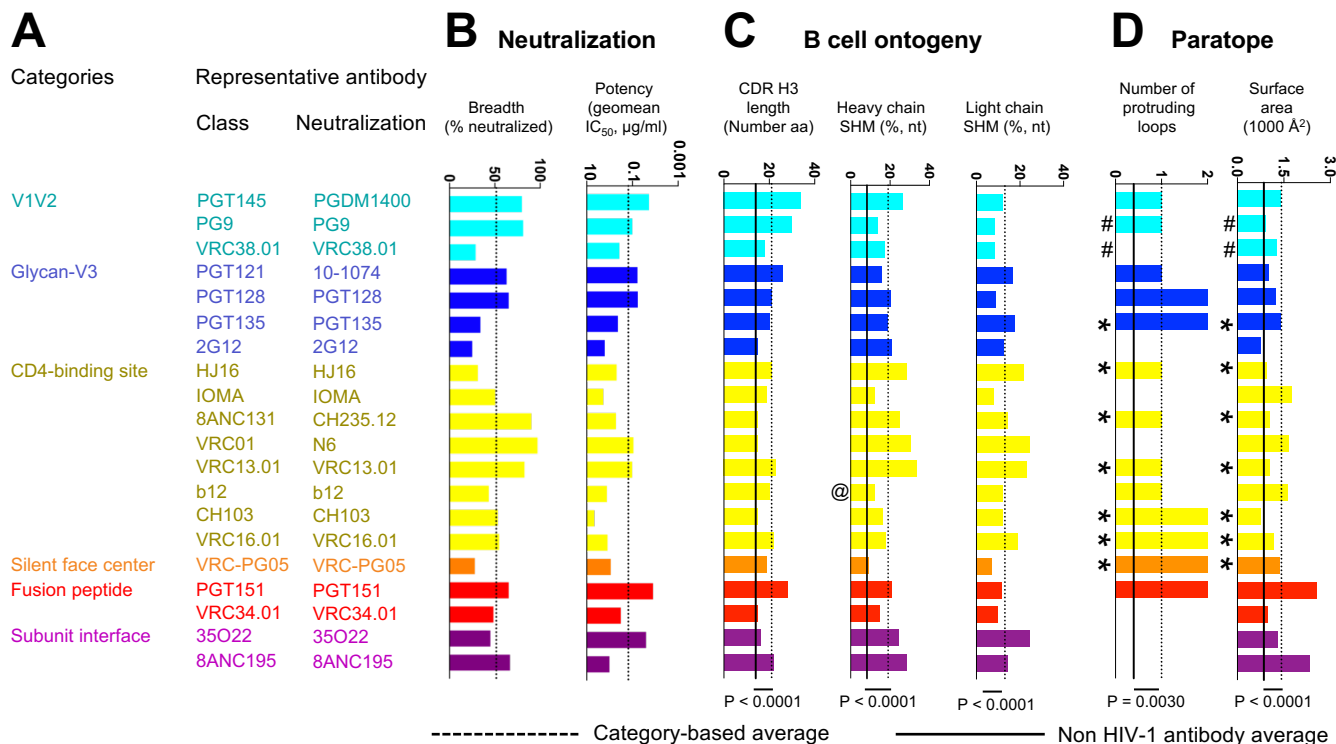
(A) Categories, classes, and representative structures. (B) Recombination properties. Bar graphs show average VH gene frequency and paratope properties of D gene, N addition and CDR H3 for each representative antibody. Average VH gene frequencies report from healthy donors in PRJNA304166. D gene and N additions occurring at the junctions are from IMGT V-quest ([http://www.imgt.org/IMGT\\_vquest/vquest](http://www.imgt.org/IMGT_vquest/vquest)). # indicates that antibody missing nucleotide sequences at junction. (C) SHM likelihood. Bar graphs show the paratope contributed from SHM



(1st amino acid to the 2nd conserved Cys) and rare SHM. @ indicates that antibody missing nucleotide sequence.

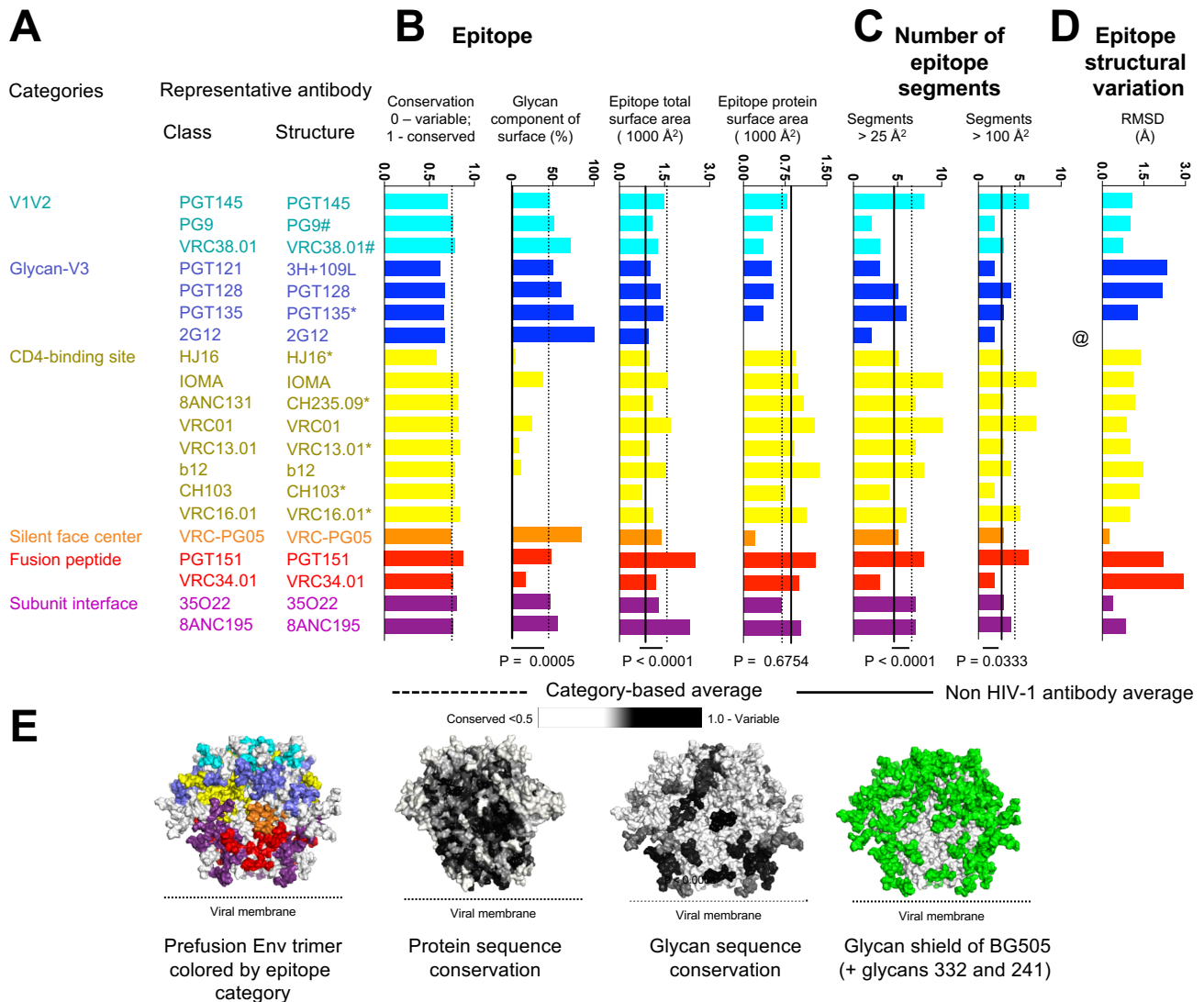


**Fig. 1. 20 classes of broadly neutralizing antibodies recognize the prefusion-closed Env trimer and segregate into 6 categories based on the Env residues with which they interact.** (A) All HIV-1 Env-antibody complexes structures in the PDB (Table S1) were assigned to classes (leftmost column, listed by the name of first reported antibody of the class) based on similarities in B cell ontogeny and mode of recognition, with a representative structure and PDB for the class (2nd and 3rd columns from left), which were chosen based on resolution and degree to which Env in the structure resembled prefusion-closed trimer; “\*” indicate structures determined in gp120-core context; “#” indicates structures determined with V1/V2 scaffold; “@” indicates glycan N295, N332, and N392 used for 2G12 epitope (see methods) (B) Prefusion-closed Env trimer with molecular surface colored by categories defined in (A). Epitope residues shared by antibodies in separate classes were colored according to surface area and requirements for antibody binding. For example, glycan N276 has interactions with antibody 8ANC195 of the subunit interface category and with multiple antibodies of the CD4-binding site category; however, because glycan N276 is required for 8ANC195, and generally only accommodated by antibodies that target the CD4-binding site, we colored glycan N276 to be part of the subunit interface category. Left image is shown with viral membrane at bottom; right image is rotated 90° to look down on the trimer apex.



**Fig. 2. Neutralization, B cell ontogeny, and paratope characteristics for 20 classes of broadly neutralizing antibodies.**

**(A)** The broadest neutralizing antibody from each class was chosen as representative. **(B)** Neutralization properties. Bar graphs show breadth and potency of each representative antibody on the 208-isolate panel, with dotted line showing averages. **(C)** Features of B cell ontogeny. CDR H3 lengths are defined based on Kabat nomenclature. @ indicates that antibody b12 is derived from a phage library. IOMA is the only antibody with CDHR H3 length under 20, SHM <12%, and breadth approaching 50%, and thus may be a promising candidate for lineage-based vaccine design. By contrast, VRC01-class antibodies have high SHM and PG9/PGT145 have unusually long CDR H3, complicating their lineage-based elicitation. **(D)** Paratope properties. “\*” indicate structures determined in gp120-core context; “#” indicates structures determined with V1/V2 scaffold. Structures denoted with “\*” or “#” were included in calculations of the category averages for protruding loop numbers, but not for paratope surface area. P-values were determined using two-tailed Mann-Whitney test.



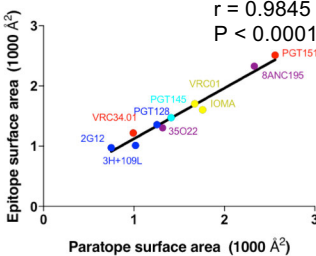
**Fig. 3. Epitope features for 20 classes of broadly neutralizing antibodies.**

(A) Categories, classes, and representative structures. “\*” indicate structures determined in gp120-core context; “#” indicates structures determined with V1/V2 scaffold; @ indicates that the epitope structural variation of antibody 2G12 was not calculated due to the lack of protein epitope. Structures denoted with “\*” or “#” were not included in calculations of the category averages for number of epitope segments and epitope total surface. Structures denoted with “\*” were not included in calculations of the category average for epitope glycan composition. (B) Epitope properties. Bar graphs show conservation and glycan-interactive surface for each representative antibody, with dotted line showing averages. “\*” indicate structures determined in gp120-core context, where glycan surface areas may not reflect trimer interactions. (C) Number of epitope segments, shown for segments with  $>25 \text{ \AA}^2$  or  $> 100 \text{ \AA}^2$  surface area. VRC34 has low number epitope segments and low glycan context (black arrow), and this may explain why fusion peptide immunization are yielding promising results from epitope-based vaccine design. (D) Observed variation in epitope structure, as calculated from RMSD of recognized amino acid residues, computed pairwise among class members of each category and unliganded trimer. (E) Prefusion-closed Env trimer colored by epitope and by various epitope features. P-values were determined using two-tailed Mann-Whitney test.

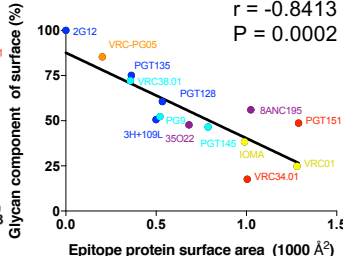
**A**

	Pearson r													
P-value	Neutralization breadth	Neutralization potency	CDR H3 length	Heavy chain SHM	Light chain SHM	Number of protruding loops	Paratope surface area	Epitope conservation	Epitope glycan component surface	Epitope total surface area	Epitope protein surface area	Number of epitope segments (>25 Å²)	Epitope structural variation	
Neutralization breadth				0.3278	0.4524	0.2276	-0.0107	0.4370	0.3998	-0.6594	0.4245	0.5902	0.5983	0.0805
Neutralization potency	0.2773		-0.3907	-0.2988	-0.2201	0.1556	-0.0290	0.0913	0.3699	-0.0504	-0.0706	-0.0138	-0.1108	
CDR H3 length	0.1707	0.0982		0.0407	-0.1448	0.3045	0.3393	-0.1152	-0.0732	0.3136	0.0610	0.1803	0.1014	
Heavy chain SHM	0.0518	0.2141	0.8685		0.7188	-0.2426	0.3106	0.0468	-0.2092	0.3643	0.4136	0.3577	-0.1581	
Light chain SHM	0.3488	0.3652	0.5542	0.0005		-0.1340	-0.0192	0.0134	-0.2637	-0.0510	0.3466	0.2205	-0.2166	
Number of protruding loops	0.9654	0.5247	0.2049	0.3170	0.5844		0.2303	-0.0794	0.3075	0.2376	-0.0563	-0.0362	0.1359	
Paratope surface area	0.2066	0.9365	0.3374	0.3823	0.9580	0.5220		0.7110	-0.2303	0.9845	0.7525	0.6884	-0.2865	
Epitope conservation	0.0899	0.7100	0.6386	0.8491	0.9565	0.7466	0.0211		-0.4444	0.6961	0.5435	0.7113	-0.2370	
Epitope glycan component surface	0.0103	0.1930	0.8035	0.4730	0.3623	0.2849	0.5221	0.1114		-0.2003	-0.8413	-0.4554	-0.4357	
Epitope total surface area	0.2214	0.8900	0.3776	0.3007	0.8887	0.5086	<0.0001	0.0254	0.5790		0.7428	0.6110	-0.2065	
Epitope protein surface area	0.0078	0.7740	0.8041	0.0784	0.1461	0.8188	0.0120	0.0162	0.0002	0.0139		0.7204	-0.2041	
Number of epitope segments (>25 Å²)	0.0676	0.9698	0.6181	0.3101	0.5403	0.9208	0.0277	0.0211	0.1860	0.0606	0.0188		0.2041	
Epitope structural variation	0.7508	0.6617	0.6888	0.5309	0.3880	0.5908	0.4548	0.3437	0.1367	0.5939	0.4165	0.0317		

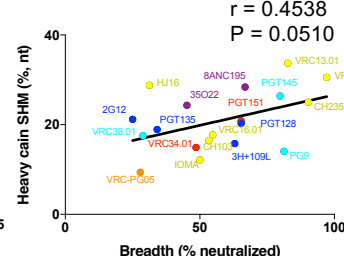
**B**



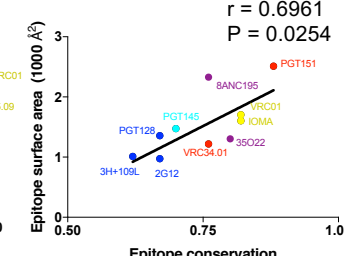
**C**



**D**

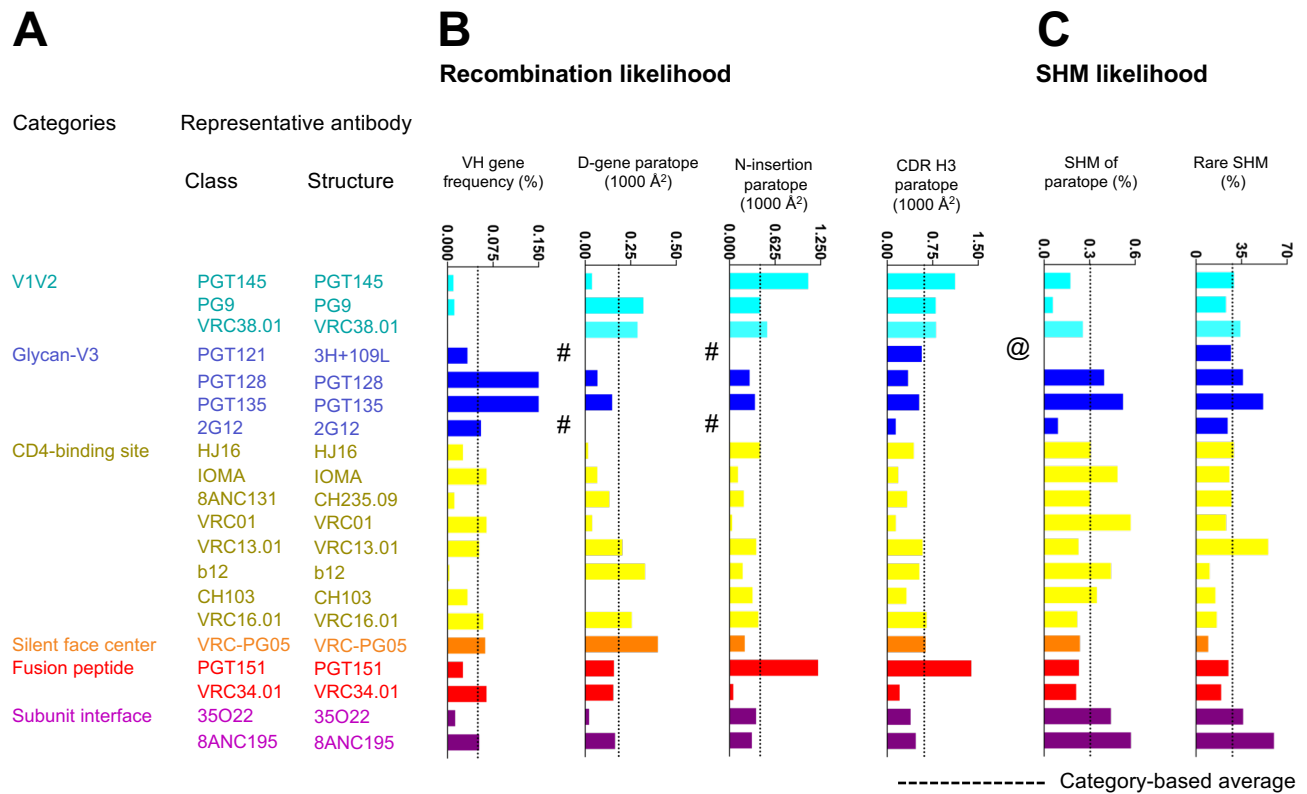


**E**



**Fig. 4. Correlations of neutralization and features of epitope and paratope reveal underlying relationships.**

Correlation matrices are shown for properties of neutralizing antibodies that recognize the prefusion closed Env trimer. **(A)** Pearson correlations ( $r$ ) are provided in the upper right; P-values (not adjusted for multiple comparisons) are provided in the lower left. Correlations with P-value of less than 0.06 were highlighted with dark red font, and correlations with P-value of less than 0.05 after Bonferroni correction were highlighted with both dark red font and orange fill. **(B)** Correlation between paratope surface area and epitope surface area. **(C)** Correlation between Epitope protein surface area and glycan component of surface. **(D)** Correlation between neutralizing breadth and heavy chain SHM. **(E)** Correlation between epitope conservation and epitope total surface area.



**Fig. 5. Factors affecting re-elicitation**

(A) Categories, classes, and representative structures. (B) Recombination properties. Bar graphs show average VH gene frequency and paratope properties of D gene, N addition and CDR H3 for each representative antibody. Average VH gene frequencies report from healthy donors in PRJNA304166. D gene and N additions occurring at the junctions are from IMGT V-quest ([http://www.imgt.org/IMGT\\_vquest/vquest](http://www.imgt.org/IMGT_vquest/vquest)). # indicates that antibody missing nucleotide sequences at junction. (C) SHM likelihood. Bar graphs show the paratope contributed from SHM (1st amino acid to the 2nd conserved Cys) and rare SHM. @ indicates that antibody missing nucleotide sequence.

**Table 1. Structural details for 20 classes of antibodies in complex with HIV-1 Env.**

Category	Class	Antibody with most informative structure	Most informative PDB	Experiment method	Resolution (Å)	Env component	Epitope Residues
V1V2	PGT145	PGT145	5V8L	Cryo-EM	4.3	Trimer	K121, T123, P124, V127, M161, T162, R166, D167, K169, I309, Q315, NAG608, T123, V127, N160, M161, T162, R166, D167, K168, K169, NAG606, NAG607, BMA608, K121, T123, P124, V127, N160, M161, T162, R166, D167, K168, K169, glycan 160
	PG9	PG9	3U2S	X-ray	1.8	Scaffolded V1V2	C157, S158, N160, T162, T163, K166, D167, R168, K169, Q170, K171, V172, N173, glycan 160, glycan 173
	VRC38.01	VRC38.01	5VGJ	X-ray	3.5	Scaffolded V1V2 region	H130, T132, N133, V134, T135, I136, N156, C157, S158, F159, N160, K168, K171, E172, Y173, glycan 133, glycan 156, glycan 160
Glycan-V3	PGT121	3H+109L	5CEZ	X-ray	3.0	Trimer	V134, T135, N136, A137, T139, D140, D321, I322, I323, G324, D325, I326, R327, Q328, H330, T415, P417, glycan 156, glycan 301, glycan 332
	PGT128	PGT128	5ACO	Cryo-EM	4.4	Trimer	V134, T135, N136, N137, I138, L175, T297, P299, N301, D321, I322, I323, G324, D325, I326, R327, V442, R444, C445, glycan 156, glycan 301, glycan 332
	PGT135	PGT135	4JM2	X-ray	3.1	gp120	P299, I323, G324, Q328, A329, H330, R335, M373, Y384, N386, Q389, T408, E409, T415, P417, C418, R419, K421, glycan 332, glycan 392, glycan 386
	2G12	2G12	1OP5	X-ray	3.0	N-glycan	Glycans 295, 332, 392 (Sanders et al. 2002 )
CD4-binding site	HJ16	HJ16	4YE4	X-ray	2.7	gp120	
	IOMA	IOMA	5T3Z	X-ray	3.5	Trimer	T278, N279, N280, A281, T283, V360, Q362, P363, S364, S365, G366, G367, I371, T455, R456, D457, G458, G459, N460, N461, N462, K463, I469, R471, P472, G473, G474, D476, glycan 276
	8ANC131	CH235.09	5F9O	X-ray	1.9	gp120	K 97, E102, T198, T257, T278, N279, N280, A281, K282, R360, S365, G366, G367, D368, E370, V371, M426, W427, Q428, I430, T455, R456, D457, G458, G459, S460, T461, N462, S463, E466, T467, R469, G472, G473, D474, M475, R476, glycan 197, glycan 276, glycan 363, K 65, H 66, R308, glycan 262, glycan 301
	VRC01	VRC01	5FYJ	X-ray	3.1	Trimer	W 96, K 97, E275, N276, L277, T278, N279, N280, A281, K282, T283, S365, G366, G367, D368, I371, M426, G429, G431, Q432, T455, R456, D457, G458, G459, A460, N461, R469, P470, G471, G472, N474, D477, R480, glycan 276
	VRC13.01	VRC13.01	4YDJ	X-ray	2.3	gp120	W 96, K 97, E275, N279, N280, A281, K282, T283, R327, P364, S365, G366, G367, D368, L369, I371, T372, M373, Y384, K419, I420, K421, I423, T455, G458, G459, R469, P470, G471, G472, G473, N474, D477, R480, glycan 386
	b12	b12	5VN8	Cryo-EM	3.6	Trimer	K178, L179, V181, V182, P183, L184, E185, N185, K185, N185, N187, I188, T189, R192, I194, N195, N197, T198, S199, T257, N280, A281, H363, S364, S365, G366, G367, D368, P369, E370, I371, V372, T373, Y384, N386, Q417, C418, R419, K421, N425, M426, W427, Q428, G429, V430, G431, T455, R456, D457, R460, R469, G471, G472, G473, N474, M475, glycan 197, glycan 362, glycan 386
	CH103	CH103	4JAN	X-ray	3.2	gp120	S256, D279, N280, A281, E362, H364, S365, G366, G367, D368, L369, E370, I371, T455, R456, D457, G458, G459, N460, D461, D462, N463, R469, P470, G471, G472
	VRC16.01	VRC16.01	4YDK	X-ray	2.1	gp120	Q105, L122, T123, G124, G198, S199, T257, E275, N280, A281, K282, T283, P364, S365, G366, G367, D368, L369, E370, I371, H375, N425, M426, W427, Q428, G429, G431, Q432, T455, R456, D457, G458, G459, A460, T463, T467, R469, P470, G471, G472, G473, N474, D477
Silent face center Fusion peptide	VRC-PG05	VRC-PG05	6BF4	X-ray	2.4	gp120	N262, K290, S291, V292, E293, D334, E337, L446, S447, N448, glycan 262, glycan 295, glycan 448
	PGT151	PGT151	5FUU	Cryo-EM	4.2	Trimer	A 58, P 79, N 80, Q 82, E 83, V 84, V 85, A224, T244, V245, Q246, glycan 241, glycan 262, glycan 295, glycan 448, A512, V513, G514, I515, G516, A517, V518, F519, G521, F522, R542, L543, S546, G547, V549, Q550, Q551, Q552, N553, N554, L592, N637, Y638, S640, E641, Y643, T644, E647, glycan 611, glycan 637
	VRC34	VRC34.01	5I8H	X-ray	4.3	Trimer	N 80, Q 82, E 83, I 84, H 85, E 87, K229, D230, K231, S241, glycan 88, A512, V513, G514, I515, G516, A517, V518, F519, L520, G521, E648
Subunit interface	35O22	35O22	5CEZ	X-ray	3.0	Trimer	G527, S528, T529, A532, R617, N618, S620, E621, D624, N625, M626, T627, L629, Q630, K633, glycan 618, E 87, N 88, T 90, E 91, E 92, P238, P240, glycan 234, glycan 88
	8ANC195	8ANC195	5CJX	X-ray	3.6	Trimer	S612, S613, W614, S615, N616, R617, M626, L629, Q630, D632, K633, E634, S636, N637, Y638, glycan 637, V 44, W 45, K 46, D 47, V 89, T 90, E 91, E 92, F 93, N 94, K 97, T236, G237, P238, E275, N276, I277, T278, K351, H352, F353, R456, K487, E492, glycan 234, glycan 276