

## Trimeric HIV-1-Env Structures Define Glycan Shields from Clades A, B and G

Guillaume B. E. Stewart-Jones<sup>1</sup>, Cinque Soto<sup>1</sup>, Thomas Lemmin<sup>1,2</sup>, Gwo-Yu Chuang<sup>1</sup>, Aliaksandr Druz<sup>1</sup>, Rui Kong<sup>1</sup>, Paul V. Thomas<sup>1</sup>, Kshitij Wagh<sup>3</sup>, Tongqing Zhou<sup>1</sup>, Anna-Janina Behrens<sup>4</sup>, Tatsiana Bylund<sup>1</sup>, Chang W. Choi<sup>1</sup>, Jack R. Davison<sup>5</sup>, Ivelin S. Georgiev<sup>1</sup>, M. Gordon Joyce<sup>1</sup>, Young Do Kwon<sup>1</sup>, Marie Pancera<sup>1</sup>, Justin Taft<sup>1</sup>, Yongping Yang<sup>1</sup>, Baoshan Zhang<sup>1</sup>, Sachin S. Shivatare<sup>6</sup>, Vidya S. Shivatare<sup>6</sup>, Chang-Chun D. Lee<sup>6</sup>, Chung-Yi Wu<sup>6</sup>, Carole A. Bewley<sup>5</sup>, Dennis R. Burton<sup>7,8</sup>, Wayne C. Koff<sup>9</sup>, Mark Connors<sup>10</sup>, Max Crispin<sup>4</sup>, Ulrich Baxa<sup>11</sup>, Bette T. Korber<sup>3</sup>, Chi-Huey Wong<sup>6,12</sup>, John R. Mascola<sup>1</sup> and Peter D. Kwong<sup>1\*</sup>

<sup>1</sup> Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA.

<sup>2</sup> Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, California 94143, USA.

<sup>3</sup> Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM 87544, USA

<sup>4</sup> Oxford Glycobiology Institute, Department of Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 3QU, United Kingdom.

<sup>5</sup> Laboratory of Bioorganic Chemistry, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland, USA.

<sup>6</sup> Genomics Research Center, Academia Sinica, 128 Academia Road, Section 2, Nankang, Taipei 115, Taiwan.

<sup>7</sup> Department of Immunology and Microbial Science, International AIDS Vaccine Initiative Neutralizing Antibody Center, Center for HIV/AIDS Vaccine Immunology and Immunogen Discovery, The Scripps Research Institute, La Jolla, CA 92037, USA.

<sup>8</sup> Ragon Institute of Massachusetts General Hospital, Massachusetts Institute of Technology and Harvard University, Boston, MA 02142, USA.

<sup>9</sup> International AIDS Vaccine Initiative, New York, NY 10038, USA.

<sup>10</sup> HIV-Specific Immunity Section, Laboratory of Immunoregulation, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA.

<sup>11</sup> Electron Microscopy Laboratory, Cancer Research Technology Program, Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, Frederick, Maryland 21702, United States of America

<sup>12</sup> Department of Chemistry, The Scripps Research Institute, La Jolla, CA 92037, USA.

\* Correspondence should be addressed to P.D.K. [E-mail: pdkwong@nih.gov](mailto:pdkwong@nih.gov)

The HIV-1-envelope (Env) trimer is covered by a glycan shield of ~90 N-linked oligosaccharides, which comprises roughly half its mass and is a key component of its evasion from antibody. To obtain an atomic-level understanding of this shield, we used an antibody lattice to crystallize fully glycosylated Env trimers from three clades at 3.4-3.7 Å resolution. These structures revealed prominent ridges of hydrogen-bonded oligomannoses, supported by N-acetylglucosamine stems and protruding 20 Å above the protein surface, to extend from the membrane-proximal base of the trimer along exposed peripheral surfaces to cover the trimer variable-loop apex. The revealed features delineated a taxonomy of N-linked glycan-glycan interactions. We observed crowded and dispersed glycans to be differently ordered, conserved, and recognized by antibody, with structures, molecular dynamics, and glycan arrays revealing a diversity in oligosaccharide affinity and a requirement for accommodating glycan amongst known broadly neutralizing antibodies that target the glycan-shielded trimer.

## INTRODUCTION

Glycan shields reside at the interface between virology, glycobiology and immunology. Very few pathogens withstand the antibody-mediated onslaught of the humoral immune system, but HIV can. Indeed, despite the presence of high titer HIV-reactive antibodies, sustained viremia is frequently observed over years of chronic HIV infection (Wei et al., 1995). The primary mechanism appears to be an evolving glycan shield (Wei et al., 2003), which couples the typical mutational rate of a small RNA virus (Jenkins et al., 2002) with an extraordinary level of N-linked glycosylation on the HIV-1 envelope (Env) glycoproteins (Allan et al., 1985), gp120 and gp41, which trimerize to form the HIV-1 viral spike, the only exposed viral antigen on the virion surface (reviewed in (Wyatt and Sodroski, 1998)). A median of 93 N-linked glycans (95% CI = 75-105, based on 2994 Env M-group sequences ([www.hiv.lanl.gov](http://www.hiv.lanl.gov))), cloak the surface of the trimeric spike and comprise over half its mass (Leonard et al., 1990) (**Table S1**).

Each N-linked glycan is encoded by a tripeptide, Asn-X-Ser/Thr, the N-glycan sequon (reviewed in (Kornfeld and Kornfeld, 1985)), which is recognized by oligosaccharyltransferase on nascent polypeptides as they extrude into the lumen of the endoplasmic reticulum to couple a fully formed Glc<sub>3</sub>Man<sub>9</sub>(GlcNAc)<sub>2</sub> precursor (2371 Da) to Asn in the N-glycan sequon (Llanover and Lennarz, 1980). The Glc<sub>3</sub>Man<sub>9</sub>(GlcNAc)<sub>2</sub> precursor is trimmed by glycosidases through high mannose forms, then modified by glycotransferases, which assemble complex glycans in the Golgi meant for signaling and other glycobiology functions utilized by nearly all secreted or cell-surface expressed proteins in humans (Varki, 2011).

Because of their ubiquitous presence, all of these forms of N-linked glycosylation high mannose through complex are privileged from antibody recognition through check-point induced tolerance of the functional antibody repertoire. While antibodies can recognize non-self glycans (such as their recognition of the glycan antigens that define the human blood groups)

(Stowell et al., 2008), if antibody recognition of a viral *N*-linked glycan were to result in cross-recognition of host *N*-linked glycans, such self-recognition results in the anergy or apoptosis of the *N*-glycan-reactive B cell (Wardemann et al., 2003). The upshot is that antibodies only recognize viral *N*-linked glycans in the context of a viral protein (a glycoprotein epitope) (Kong et al., 2013; McLellan et al., 2011; Pejchal et al., 2011) or in the context of at least two other viral glycans (an exclusively glycan-based epitope) (Sanders et al., 2002; Scanlan et al., 2002).

The ability of glycan shields to thwart the humoral response is demonstrated in part by their ubiquity: viruses as diverse as hepatitis C virus (HCV; *Flaviviridae*) (Helle et al., 2011), Epstein Barr virus (EBV; *Herpesviridae*) (Szakonyi et al., 2006), and Lassa virus (LASV; *Arenaviridae*) (Sommerstein et al., 2015), along with the aforementioned HIV-1, have extensive *N*-linked glycosylation, which covers exposed protein surfaces and whose glycan mass can exceed that of the protein component. Of these, the HIV-1 glycan shield is perhaps the most well-characterized. Structural studies of the HIV-1 Env trimer or its gp120 subunit have revealed tantalizing glimpses; however, glycosylated gp120 structures lack the interprotomer interactions of the pre-fusion closed state (Diskin et al., 2013; Kong et al., 2015), and at best only about a quarter of the glycan content has been observed with published trimeric structures (Garces et al., 2015; Julien et al., 2013; Kwon et al., 2015; Lee et al., 2015; Lee et al., 2016; Lyumkis et al., 2013; Pancera et al., 2014; Scharf et al., 2015) (**Table S2**). Thus, despite its importance, knowledge of the glycan shield has been primarily antibody-epitope focused, with little understanding of its cumulative characteristics, aggregate properties, and overall atomic-level structure. Here we present the 3.4 Å structure of a fully glycosylated HIV-1 pre-fusion trimer from clade G and characterize the organization of glycans and clustering of oligomannose residues. Additionally we present the 3.7 Å structures of fully glycosylated HIV-1 pre-fusion

trimers from clades A and B and carried out molecular dynamics simulations and other analyses to provide insight into antibody-glycan interactions. The results reveal the HIV-1 glycan shield to form a network of interlocking oligosaccharides, substantially ordered by glycan crowding, which encase the protein component of the trimer and enable HIV-1 to avoid most antibody-mediated neutralization.

## **RESULTS**

### **Design and Production of Diverse Pre-Fusion HIV-1 Trimers**

To provide a general understanding of the HIV-1 glycan shield, we sought to determine crystal structures of glycan-shielded Env trimers from diverse HIV-1 clades. The production of pre-fusion closed trimers from diverse HIV-1 strains has been an obstacle to structural and immunogenicity studies. Sanders, Moore and colleagues identified a clade A strain (BG505), stabilized by a disulfide (SOS) and isoleucine to proline mutation (IP) and truncated at residue 664, which formed stable soluble trimers (Sanders et al., 2013). However, when SOSIP mutations were incorporated into other HIV-1 strains, many misfolded or formed aberrant structures, although more recently a few additional well-folded HIV-1 Env strains in the SOSIP context have been identified (de Taeye et al., 2015; Julien et al., 2015; Pugach et al., 2015). As the addition of antibody VRC01 to BaL gp160 HIV-1 virions stabilizes the pre-fusion closed state as assessed by cryo-electron tomography (Tran et al., 2012), we investigated whether the co-expression of antibody VRC01 with HIV-1 Env trimers might promote formation and stabilization of pre-fusion closed trimers.

To aid in VRC01-trimer complex formation, we designed cysteine mutations across the interface of the heavy chain of antibody VRC01 (A60C<sub>HC</sub>) and the CD4-binding site of HIV-1

gp120 (G459C<sub>gp120</sub>) from a range of HIV-1 clades (**Figure S1A**) (for clarity, the molecule is added as subscript, see Extended Experimental Procedures). We succeeded in expressing covalent complexes of scFv or Fab VRC01 with SOSIP-stabilized trimers (**Figure S1B-J**), as stable pre-fusion closed Env trimers from diverse HIV-1 strains.

### Structure of a Fully Glycosylated HIV-1 Env Trimer from Clade G

We observed the lattice formed by antibodies PGT122 and 35O22 in the BG505 SOSIP.664 T332N trimer structure (PDB 4TVP) (Pancera et al., 2014) to be dependent on antibody-antibody contacts (**Figure 1A**), and complexed Fabs PGT122 and 35O22 with fully glycosylated clade G X1 193.c1 SOSIP.665 G459C trimers covalently linked with scFv VRC01 (A60C<sub>HC</sub>) (**Figure S1K**) and crystallized in hanging droplets. Data extended to a nominal 3.4 Å resolution, and the structure were refined to an  $R_{work}/R_{free}$  of 21.4%/27.2% with an average of 4.8 refined saccharide units/sequence and a solvent accessible-surface area for protein versus carbohydrate of 74,915 Å<sup>2</sup> and 73,948 Å<sup>2</sup> respectively. (**Tables S3 and S4**, Methods and Extended Experimental Procedures). Electron density was observed for all HIV-1 Env residues between 31 and 665 of strain X1 193.c1, except for 3 gp120 C-terminal amino acids of the engineered R6 furin cleavage site from 508-510. All five variable loops (V1-V5) of gp120 were resolved along with the complete fusion peptide of gp41, which was partially stabilized by an interprotomer interaction with a neighboring gp41 (**Figure S2A**). In terms of glycan, ultra-performance liquid chromatography and mass spectrometry confirmed the *N*-linked glycan to be primarily high mannose, with, Man<sub>5</sub>GlcNAc<sub>2</sub> (Man-5), constituting 31.4% of the glycan, and Man-6 and Man-7 being much less predominant than Man-8 and Man-9 (**Table S5**). We observed electron density for 29 of the 31 *N*-linked glycans of X1 193.c1 strain, which were

distributed across the trimer protein surface (**Figures 1B, S1L, S1M**). 19 of these glycans were largely exposed to the solvent space between individual trimers suspended by the PGT 122 and 35O22 Fabs in the crystal lattice (**Figures S1N, S1O**). Three glycans showed extensive interactions with antibody including N88 to 35O22, N276 to VRC01, and N332 to PGT122, whereas no glycans were involved in forming crystallization contacts, thereby providing a mostly unperturbed view of the HIV-1 trimeric glycan shield.

We observed numerous occurrences of glycans organized into discrete structural groups. For example, at the apex of the trimer where the three C-strands of V1/V2 associate, glycan N160 was substantially ordered: two N-acetyl glucosamines (GlcNAc<sub>2</sub>) and the -mannose sugar (Man ) projected perpendicularly 20 Å from the protein surface (**Figure 1B, right insert**), with oligomannose branches engaged horizontally, making direct contacts with neighboring N1 60 oligomannose branches and assembling into an interprotomeric oligomannose ring about ~1 5 Å in diameter. The N1 60 glycans were buttressed by V1 side chains Glu1 87A and Lys1 87B, which contacted N1 60 mannose and N-acetylglucosamine residues, respectively. Additionally, mannose residues of the V1 glycan N188 contacted mannose residues of the N160 glycan ring. The conformation of the N160-glycan ring thus appeared to depend on the V1 sequence, its structure, and its glycosylation.

The N-glycan density on the gp120 outer domain was particularly high (**Figure 1B, left insert**). This region contained glycans that were most clearly defined in the electron density (except for glycans bound by antibody). The structure revealed 1 ,4GlcNAc stems to extend perpendicularly ~20 Å from the protein surface, with oligosaccharide branches forming a lateral canopy of glycans that extended from the membrane-proximal 35O22 epitope all the way to the variable loop cover of the trimer apex. In this ridge of interconnected glycans,

oligosaccharides showed extensive electron density and allowed five Man-8 and Man-9 glycans to be modeled. However in other regions such as the glycans from gp41, only 2-3 sugar residues could be fitted to the electron density. In general, glycans not bound by antibody had *B*-factors similar to those observed for hypervariable loop residues (**Figure S2B, C**).

### **Structures of Fully Glycosylated HIV-1 Env Trimers from Clades A and B**

We used the PGT122-35O22-HIV-1 trimer lattice to crystallize a fully glycosylated VRC01-complexed Env trimer from the clade B strain JR-FL as well as a fully glycosylated Env trimer from the clade A strain BG505 (**Figure S2D**). Diffraction from these crystals extended to nominal resolutions of 3.7 Å, and molecular replacement and refinement led to  $R_{work}/R_{free}$  of 25.3%/30.9% and 26.0%/30.7% for JR-FL and BG505, respectively, with an average of 4.6 and 3.8 refined saccharide units/sequon respectively and protein versus carbohydrate solvent accessible-surface area values of 66,575 versus 68,502 Å<sup>2</sup> (JR-FL) and 67,960 versus 60,640 Å<sup>2</sup> (BG505) (**Tables S1 and S3**). Interestingly, despite general similarity of the lattice, the crystallizing antibodies did show strain-dependent variation, especially with N137, a glycan known to be important in PGT122 recognition (Garces et al., 2015; Garces et al., 2014). N137 was disordered in JR-FL, ordered in BG505 and displaced by a V1 extension in X11 91 .c 1 (**Figure S2E**).

Glycan compositions were high mannose, consistent with the GnTI-deficient cell line used for production, with Man-5 constituting 24.6% and 15.3% of the glycan from JR-FL and BG505 strains, respectively, and with Man-8 and Man-9 dominating over Man-7 and Man-6 (**Table S5, Figure S3A**). The glycan shields of BG505 SOSIP.664 and JR-FL SOSIP.664 revealed varied organizations of glycans on the HIV-1 trimer surfaces (**Figures 2A, B, S2F**),



likely arising from differences in the positions of glycan sequons as well as differences in Env loop lengths and amino acids near N-linked glycans. Stem structures involving Man 1-4GlcNAc 1,4GlcNAc, however, were generally conserved and projected similarly from the protein surface (**Figure 2C**), with oligomannose arms forming strain-dependent organizations. Overall, each of the three glycan structures appeared to provide substantial shielding of the pre-fusion HIV-1 Env trimer (**Figure 2D**).

The protein structures of the clades A, B and G Env analyzed here serve to define a conserved pre-fusion HIV-1 Env core (**Figure 2E, F**), with structural variability for both surface residues and internal residues such as the interhelical region of gp41 (**Figure S2A, D**). Comparison of the glycan order present in the three gp140 strains showed high correlation between different clades ( $r = 0.766 - 0.798$ ) (**Figure 2G**). Thus, the clade A and clade B Env structures confirmed much of the ordered glycan shield observed with the clade G Env structure, and along with the clade G structure defined conserved glycan and protein core structures, as well as variable features.

### **A Distance-Based Classification of N-Linked Glycan-Glycan Interactions**

Analysis of the inter-sequon neighbor distances revealed glycans on the HIV-1 trimers to have average nearest inter-glycan sequon distance peaking at 12 Å, and we classified these into three categories: (I) 4-7 Å, (II) 9-18 Å and (III) >20 Å (as measured by glycan sequon Asn 2-2 atom distance) (**Figure 3A**). The dispersal of HIV-1 Env glycans was distinct from that observed with other type 1 fusion machines (**Figure 3B**). With HIV-1, a variety of glycan-glycan interactions was observed (**Figures 3C, S3B-E**). In category I, glycans made contacts primarily via GlcNAc stem residues. When glycan sequons were closely juxtaposed (e.g. 4 Å apart), we

observed glycan forking (**Figure 3C upper left**); however, when the distance between sequons was slightly increased (e.g. 6 Å apart), glycan stems were observed to align in a parallel manner (**Figure 3C lower left**). Category II glycans with 9–18 Å inter-sequon distances lacked interactions between GlcNAc<sub>2</sub> stems; rather GlcNAc<sub>2</sub> stems were generally solvated, with glycan-glycan interactions occurring via oligomannose branches. These generally contacted other oligomannoses (branch-branch engagement), although interactions with GlcNAc<sub>2</sub> stems (stem-branch association) were also observed (**Figure 3C, middle-left panels**). More complex associations or network of associations (**Figure S3B-D**) were observed comprising a variety of interlocking assemblies with two and sometimes three glycans clustering together (**Figure 3C middle right panels**), and including some long-range inter-glycan interactions, where sequons can be separated by 30 Å (**Figure 3C, lower right**). Despite the ability of glycans to form long-range interactions, these were not observed with category III glycans. Rather, category III glycans showed electron density for only two or three protein-proximal saccharide units and lacked clearly defined glycan-glycan interactions (**Figure 3C, upper right**).

The various observed glycan-glycan interactions may inhibit steric access to the protein surface. Steric inhibition might occur through a ‘local effect’, where a glycan shields a local patch (such as with glycan N197, **Figure 3C, upper right**), through a ‘mesh effect’, where two or more oligosaccharides interact locally (such as with X1193.c1 glycans N241 and N293, **Figure 3C, lower right**), or through a substantially interacting glycan shield by creating an oligosaccharide canopy over the protein surface which may impede protein surface access (such as the glycan-shielded apex of X1193.c1, **Figure 3C, upper middle right**). In addition to these ordered effects, the disordered regions of

the glycan shield may form entropic or dynamic barriers, also inhibiting access to the protein surface of Env.

### Surface Glycan Density and Glycan Order

To understand the degree of ordered glycans on these fully glycosylated Env trimers, we investigated parameters that correlated with increased glycan order. The correlation between the number of ordered saccharide units in a glycan versus its nearest neighbor distance was not significant (**Figure S3F-H**), however, we did observe strong correlations between the number of ordered saccharide units and the number of surrounding active sequons within spherical radii ranging from 10 to 55 Å (**Figure 4A**). When considering all three Env trimer structures, the Pearson correlation coefficient between the number of ordered saccharide units and the number of active sequons became significant at 20 Å, which may relate to a first shell of supporting glycans, and peaks at a radius of 50 Å, which may include a second shell of supporting glycan (**Figure 4B**). Application of this radial cutoff to each Env trimer structure yielded similar correlations (**Figure S4A**) and allowed for the definition of two distinct classes of glycans:

crowded, based on the number of neighboring sequons and the number of ordered saccharide units. By using a 2 x 2 contingency table, we determined 15 neighboring sequons to provide the optimal separation between the two classes (see Methods for details and **Figure S4B, C**). Thus, a glycan was classified as being crowded if it had 15 or more neighboring glycans and dispersed if it had fewer than 15 neighboring glycans, within a 50 Å radius (**Figures 4C, S4D**).

To visualize differences between the location of crowded and dispersed glycans, we generated protein surface representations, which highlighted the 10 Å vicinity around each

glycan sequon. This revealed that crowded glycans generally occupied the membrane-distal region of the Env trimer as well as the outer ridge of each gp120 (especially the outer domain of gp120), whereas dispersed glycans were generally located in the membrane-proximal region of gp41 (**Figure 4D, upper row**). A common feature shared between all three structures was the presence of glycan-free surface (e.g. surface more than 10 Å from any glycan sequon), mostly located within the recessed regions between protruding outer domains. Strong concordance was observed between the three trimer structures for surfaces associated with crowded and dispersed glycans and the number of saccharide units on each sequon, except that sequons in Env variable loops occasionally contained only a few ordered saccharide units, despite being surrounded by (**Figure 4D, lower row**). Moreover, analysis of M-group Envs showed the pattern of glycan crowding and dispersal observed in the BG505, JR-FL and X1 193.c1 trimer structures to be representative of M-group glycans (**Figure 4E, Figure S4E**).

### **Molecular Features Associated with Crowded and Dispersed Glycans**

To provide additional insight into the biological characteristics of crowded and dispersed glycans, we analyzed the occurrence of crowded and dispersed glycan (and their 10 Å associated surfaces) within the context of the 3 -propellar blades of the prefusion closed trimer (**Figure 5A**). Notably, crowded glycans and their 10 Å associated surfaces were located primarily on convex surfaces of the trimer and covered the entire membrane-distal apex and glycan-V3 region (**Figure 5A, middle left**). Half-way down the trimer, the peripheral surfaces of the middle of the trimer were mostly covered by crowded glycan, whereas surfaces closer to the trimer axis were generally not glycosylated (that is, over 10 Å from the nearest glycan sequon) (**Figure 5A,**

**middle right**). The membrane-proximal surface of the trimer, and half the diameter of the trimer middle, was equally covered by all three types of surfaces (**Figure 5A, far right**).

As dense occurrences of glycan are known to inhibit glycan processing (Guttman et al., 2014; Pritchard et al., 2015), we would expect glycan processing to be greater with dispersed glycan. We used the glycan-type identified on BG505 SOSIP.664 (Guttman et al., 2014) and observed the number of glycan processing steps to be indeed significantly higher for dispersed versus crowded glycan ( $p = 0.0002$ , Mann Whitney test) (**Figure 5B, upper left**). We also observed higher degree of glycan sequon conservation for dispersed glycans compared to crowded glycans ( $p = 0.0245$ , Mann Whitney test), using the glycan types defined from M-group analysis. (**Figure 5B, lower left**). Crowded glycans were significantly enriched ( $p = 0.014$ , Mann Whitney test) for glycans showing significant frequency difference between M-group subtypes (27 out of 35), while dispersed glycans were mostly conserved across subtypes (**Figure S5A**). We also assessed surface residues within 10 Å of crowded or dispersed glycans based on M-group analysis and observed surface-amino acids to be more variable near glycans (**Figure 5B, upper right**). Broadly neutralizing antibodies showed an increased frequency of recognition near N-linked glycans, with a further increase for regions of crowded glycans (**Figure 5B, lower right**). Thus crowded and dispersed glycans are differently ordered, processed and conserved, with glycosylated regions most frequently recognized by broadly neutralizing antibodies.

### **Insight from Molecular Dynamics: Known Broadly Neutralizing Antibody Need To Accommodate Glycan**

To provide additional context for crowded and dispersed glycans, we used the BG505 SOSIP.664 T332N trimer structure to carry out three 500 ns molecular dynamics simulations

with either Man-5, Man-7 or Man-9 glycans at every sequon (Movie S1). One distinct feature shared among the glycans was an extended conformation for the Man-GlcNAc-GlcNAc core at the protein proximal stem. In the crystal structures, this angle was similar for crowded and dispersed glycans, both peaking at  $145^\circ$  and similar to the angle observed from the molecular dynamics simulations (Figures 5C, S5B-D, S5F-G). The crowded and dispersed groups displayed substantially different profiles with respect to non-covalent inter-glycan contacts. In all three simulations, dispersed glycans showed a relatively lower number of glycan-glycan contacts, similar to observations in the crystal structures (Figure S5E-G). Altogether, these results indicate a strong concordance between crystallographic and MD results.

We analyzed the three 500 ns molecular dynamics simulations for glycan overlap of the volumes of bound broadly neutralizing antibodies (Figure 6A, B, Figure S6A-D and Table S6). Quantitative analysis of the extent of overlap between the volume occupied by broadly neutralizing antibodies and individual glycans throughout the trajectories (Figure 6B) showed concordance between the three oligomannose simulations. This analysis also revealed that CD4 and all of the 11 broadly neutralizing antibodies analyzed, including those targeting the CD4-binding site, showed substantial overlap with at least one glycan. The simulation confirmed that antibodies known to require select glycans for recognition (PG9, PGT122, PGT135, PGT151, 8ANC195 and 35O22) did, indeed, have these glycans occupying their bound volumes. CD4-binding site directed antibodies (b12, CH103, HJ16, VRC01, and VRC13) also showed substantial occupation of their binding volume by up to 9 glycans (Figure 6B, C). Moreover, atomic-level analysis of the frequency of overlap for required glycans and non-required glycans revealed similar distributions of glycan-antibody overlap (Figure 6D). These results indicated that known broadly neutralizing antibodies likely need to accommodate at least one glycan.

## **Insight from Glycan Arrays: Broadly Neutralizing Antibodies Display Sporadic Affinity for Oligosaccharide**

Antibodies can recognize glycans in a variety of ways, including high affinity binding of oligosaccharides. We assessed oligosaccharide binding for the 11 antibodies analyzed in Figure 6B as well as for the 16 CD4-binding site antibodies analyzed previously in the context of the CD4 supersite (Zhou et al., 2015) using an array of 40 oligosaccharides (**Figure 7A, B**) containing glycans representative of oligomannose, hybrid and complex-type glycans (**Figure S6E**).

Substantial binding was observed for antibodies PGT122 and 35O22 to glycans on the array. While both recognized oligomannose and complex-type glycans (**Figure S6F, G**), 35O22 binding appeared to be limited to sialylated complex-type glycans whereas PGT 122 recognized both, a preference not entirely explained by electrostatic interactions (**Figure S6F-I**) although a sialic-acid binding site was previously defined for PGT121 (Mouquet et al., 2012) (in this structure, the bound glycan emanates from a neighboring PGT121 Fab in the crystal lattice). Other antibodies (PG9 and PGT151) known to require glycan for recognition displayed interactions with glycans on the array; of the CD4-binding site antibodies, only VRC 13 showed significant binding to the glycan array (**Figure 7A, B**). These results indicated that, while some 11IV-1 neutralizing antibodies did have direct high affinity for oligosaccharides, many broadly neutralizing 11IV-1 antibodies did not have detectable affinity for oligosaccharide in the context of a glycan array. For example, we observed no binding signal for PGT135, which is known to have substantial interaction with N-glycans 332 and 392 in its glycopeptide epitope (Kong et al., 2013).

## Antibody Neutralization and Low Affinity Glycan Interactions

In contrast to the lack of glycan array binding by antibody VRC0 1 (or by any of the 12 antibodies from the VU-gene specific VRC01 and 8ANC131 classes) (Zhou et al., 2015) (**Figure 7B**), molecular dynamics simulations indicated the bound volume of VRC01 to have substantial overlap with glycans N197, N276, N363 and N462 (**Figure 7F**). Examination of the fully glycosylated crystal structures of scFv of VRC0 1 with pre-fusion closed clade B and clade G Envs, indicated little interaction with glycans N197, N363 or N462 (**Figure S7A, B**), but substantial interactions with glycan N276, which was highly ordered (**Figures 7C, left, S7A, B**). Similar ordering was previously observed between 45-46m2 and glycan N276 with Env in its receptor-bound conformation (Diskin et al., 2013), indicating interaction with glycan N276 in both receptor-bound and pre-fusion closed conformations.

We compared the orientations of glycan N276 from the crystal structure of glycosylated BG505 SOSIP.664 T332N, with glycan N276 from the VRC01-bound structures of JR-FL SOSIP.664 and X1 193.c1 SOSIP.665 trimers (**Figure 7C, middle and right**). The VRC01-bound structures revealed the N276 glycan to be orientated in a manner compatible with VRC01 antibody access to the CD4-binding site, allowing extensive contacts of the heavy chain with the protein surface of the CD4-binding site. Additionally, comparison of BG505 SOSIP.664 T332N and X1 193.c1 SOSIP.665 structures revealed glycan N234 to differ in angle by approximately 49°, apparently re-orientating through a domino effect with glycan N276 in a concerted mechanism with VRC01 binding (**Figure 7F, middle**).

To determine experimentally whether VRC01 could bind Man<sub>9</sub>GlcNAc<sub>2</sub>Asn, we used saturation transfer difference (STD) NMR, an NMR method that can detect protein-ligand



interactions with affinities ranging from nM to mM  $K_{\text{ds}}$  (McLellan et al., 2011). Signals belonging to the N-acetyl groups of the core GlcNAc residues yielded observable STD enhancement arising from GlcNAc and mannose ring protons (Figure 7D), consistent with VRC01 binding to the conserved GlcNAc-mannose core. No STD enhancements were observed in control spectra with a sample containing glycan only. Despite this weak affinity, mutation at VRC01 light chain residues at key contact sites with glycan N276 gave rise to a loss of neutralization (Figure 7E). This was also observed in 3BNC1 17, another VH 1-2-derived antibody that has the same light chain somatic changes as the VRC01-glycan N276 contact residues and other VRC01-like antibodies (Figure S7C-E) (Zhou et al., 2013). These results suggest light chains from the VRC01 class to have somatically matured to functionally engage glycan N276 directly. Even though the glycan affinity of antibody VRC01 was too weak to be detected with glycan arrays, the VRC01 glycan interactions were nonetheless critical for its neutralization of HIV-1. Thus, while all known broadly neutralizing antibodies that target the shielded pre-fusion closed trimer need to accommodate glycan, glycan-array interactions are often not observed (Figures 6B, S7F-H) as antibodies can use a broad range of oligosaccharide affinities to fulfill this requirement.

## DISCUSSION

Atomic-level details of glycan shields have long been sought to understand how viruses use glycan to evade host humoral immunity. Here we provide proof-of-principle for the ability of the PGT122-35O22 lattice (Pancera et al., 2014) to visualize fully glycosylated Env trimers from diverse clades. It should be possible to use this lattice or the 3H+109L-3 5O22 lattice (Garces et al., 2015) to examine different forms of N-linked glycan (e.g. on trimers expressed and processed

from T cells) or different types of conformational stabilization (e.g. SOSIP.v1-v4 or DS-SOSIP) (de Taeye et al., 2015; Kwon et al., 2015). Indeed, we have grown crystals of both clade C and 293T-derived BG505 SOSIP with complex sugars (see Extended Experimental Methods), although these crystals need further optimization for atomic-level analysis.

The visualized glycan shields from 11IV-1 clades A, B and G revealed a diversity of glycan conformations and glycan-glycan interactions (Figures 1-3, S3B-E). A conserved structural motif involved the projection of protein-proximal GlcNAc<sub>2</sub> perpendicular to the protein surface, while protein-distal mannose branches assembled in a variety of ways to form extended canopies of interlocking oligosaccharide arms approximately 20 Å from the protein surface. These glycan orientations differed from N-glycans ordered through interaction with the protein surface (such as glycan N262, in which the GlcNAc stem associates with the protein surface) (Kong et al., 2015) or from glycans bound by antibody (Kong et al., 2013; McLellan et al., 2011; Pejchal et al., 2011), which interrupt branch interactions that provide the dominant form of glycan-glycan contact (Figure 3C). In general, we would expect the overall glycan orientation and glycan-glycan interactions that we observed here to be representative of dominant low energy ground states, in the same way that crystallographically visualized side chains define amino-acid stereochemistry (Engh R. A., 1991). Perhaps relevant to this, visualized glycans were often ordered to the same extent as surface amino-acid side chains (Figure S2B, C).

Although the surface of 11IV-1 Env is perhaps the most genetically diverse region of 11IV-1, we observed a number of conserved glycan features between the pre-fusion trimer structures from clades A, B and G. In particular, we observed similarities in regions of high and low glycan crowding, which appeared general to all of group M (Figure 4E) and found regions

of high glycan crowding to be correlated with lower glycan conservation, reduced glycan processing, higher Env sequence entropy, and greater serological prevalence of broadly neutralizing antibodies (**Figure 5B**). It remains to be seen if the biological characteristics of crowded and dispersed glycan that we observed with 11IV-1 will be conserved in the glycan shields of other viruses.

The combination of structures and molecular dynamics simulations carried out here suggested known broadly neutralizing antibodies that target the pre-fusion closed trimer need to accommodate glycan (**Figure 6**). This glycan requirement may explain the observation that high autologous neutralizing titers against 11IV-1 Env trimers stabilized in the pre-fusion closed state require more than 15 weeks to develop (Sanders et al., 2015), more than 3-times the time required to develop high autologous neutralizing titers against pre-fusion trimers with less glycosylation such as influenza A hemagglutinin. Perhaps relevant to this, tier 2 virus neutralizing antibodies were recently elicited via a naturally occurring N197 glycan-deficient region of a JR-FL immunogen (Crooks et al., 2015). Moreover, removal of specific glycans has been observed to be critical for interaction of Env with germline versions of some broadly neutralizing antibodies (Correia et al., 2014; Jardine et al., 2013; Jardine et al., 2015). Selective removal of glycans obstructing sites of viral vulnerability may thus allow better access to and priming of naïve B-cell receptors, which could be matured by boosting with glycan-reverted immunogens (Dosenovic et al., 2015; Jardine et al., 2013; Jardine et al., 2015). The diversity that we observed in oligosaccharide recognition by broadly neutralizing antibodies (**Figure 7**), however, cautions against relying on any single glycan-modifying strategy as a general means of eliciting broadly neutralizing antibodies against 11IV-1; thus it may be that the presence of particular glycans such as a Man-5 glycan at residue 160 (Doria-Rose et al., 2014) or a high

mannose glycan at residue 332 (Garces et al., 2015; Garces et al., 2014; Kong et al., 2013) is critical for initiating select B cell lineages. Finally, we note that the knowledge of structurally conserved and variable protein domains from the structures of Env from three diverse clades may help to guide the engineering of genotypically diverse HIV-1 Env trimers, perhaps as artificial chimeric vaccine constructs (Gorman et al., 2016). The structures of three fully glycosylated Env trimers described here thus not only allow for a definition of the glycan shield, but also provide insight into ways the shield might be modified to enhance the elicitation of neutralizing antibody.

## **EXPERIMENTAL PROCEDURES**

### **HIV-1 gp140 trimer and antibody production**

HIV-1 gp140 proteins and antibody complexes were expressed in GnTI<sup>-</sup> cells and purified as described previously (Pancera 2014) or by using Ni<sup>2+</sup> and Strep-tag affinity chromatography followed by gel filtration. PGT122 and 35O22 Fab regions were prepared and purified via size exclusion chromatography (Supplemental Experimental Procedures).

### **HIV-1 Env trimer-Fab complex preparation and crystallization**

Purified trimers were mixed with purified PGT122 and 35O22 antibody Fabs in a 1 :2:2 molar ratio (gp140 protomer:Fab:Fab) and incubated overnight at room temperature. Complexes were further purified by size exclusion chromatography and concentrated to 6-10 mg/ml for crystallization. (Supplemental Experimental Procedures).

### **X-ray data collection, structure solution and model building**

Initial structures were solved by molecular replacement using PDB ID 4TVP coordinates as a starting model (Pancera et al., 2014). Model re-building and iterative refinement procedures were carried out as described in Supplemental Experimental Procedures and data shown in Tables S1, S3).

### **Glycan analysis by ultra performance liquid chromatography**

Glycosylation profiles of N-linked glycans of purified Env proteins and antibodies were determined by hydrophilic interaction liquid chromatography-ultra performance liquid

chromatography (HILIC-UPLC) as previously described (Pritchard et al., 2015) (Supplemental Experimental Procedures and Table S5).

### **Saturation-transfer difference nuclear magnetic resonance**

Binding of VRC01 to Man<sub>9</sub>GlcNAc<sub>2</sub>Asn was detected by NMR using saturation transfer methods as described (McLellan et al., 2011). Samples were prepared in 20 mM sodium phosphate buffer containing 50 mM sodium chloride at pH 6.8. All experiments were carried out at 298 K (Supplemental Experimental Procedures).

### **Molecular dynamics simulations and glycan-antibody overlap analysis**

All molecular dynamics (MD) simulations for the three fully glycosylated BG505 SOSIP.664 trimer models (Man-5, Man-7 and Man-9) were performed over 500 ns in explicit solvent and after superimposition of CD4 and 11 antibody structures, all glycan atoms from the MD trajectory within 3.0 Å of the antibody structure were recorded (Supplemental Experimental Procedures and Table S6).

### **Bioinformatics and definition of crowded and dispersed glycans**

Using Fishers Exact test based on a  $2 \times 2$  contingency table, groups were defined as crowded or dispersed depending on the number of neighboring sequons within a spherical shell of 50 Å and the number of saccharide units. 2994 M-group Env sequences from the from the Los Alamos HIV database ([www.hiv.lanl.gov](http://www.hiv.lanl.gov)) were threaded onto each structure and the number of neighboring active glycan sequons for each glycan was then determined and was compared to that found for BG505, JR-FL or X1 193.c1 (Supplemental Experimental Procedures).

### **Fabrication of NHS-glycan microarray and antibody-glycan binding assessment**

The 40 glycans used in the array analysis were produced by the modular synthesis method and detailed in the Supplemental Experimental Procedures.

### **Neutralization assays**

Single round of replication BG505, JR-FL and X1 193.c1 and T278A variant pseudoviruses were prepared, titrated, and used to infect TZM-bl target cells as described in the Supplemental Experimental Procedures.

### **ACCESSION NUMBERS**

Coordinates and structure factors for glycosylated X1 193.c1 SOSIP-scFv VRC01-PGT122-35O22, JR-FL SOSIP-scFv VRC01-PGT122-35O22 and BG505 SOSIP-PGT122-35O22 complexes have been deposited with the Protein Data Bank under accession codes 5FYJ, 5FYK and 5FYL respectively.

### **SUPPLEMENTAL INFORMATION**

Supplemental Information includes Extended Experimental Procedures, seven figures, and five tables and can be found with this article online at <http://dx.doi.org/xxx>.

### **ACKNOWLEDGEMENTS**

We thank \_\_\_\_\_ members of the Structural Biology Section and Structural Bioinformatics Core, Vaccine Research Center

for discussions and comments on the manuscript. Support for this work was provided by the Intramural Research Program of the Vaccine Research Center, National Institute of Allergy and Infectious Diseases (NIAID) and the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), Antibody Consortium. Additional acknowledgments and support are provided in Supplemental Information.

## **AUTHOR CONTRIBUTIONS**

GSJ, CS and PDK conceived, designed and coordinated the study; GSJ, CS, KW, GYC, TL, PDK wrote and revised manuscript and figures; JRM provided intellectual expertise and guidance; HIV-1 Env designs: GSJ, IG; Antibody-antigen covalent bond design: GSJ; Protein production: GSJ, AD, PVT, BZ; Crystal structures: GSJ, TZ, PVT, MP, PDK; Molecular Dynamics: GSJ, CS, TL, PDK; Structural analysis: GSJ, CS, KW, TL, PDK; Negative stain EM: UB; VRC01 mutagenesis: GSJ, MGJ; Neutralization assays: RK, JT, CWC, JRM; Glycan UPLC analysis: AJB, M Crispin; Glycan arrays: TZ, SSS, VSS, CCL, CYW, CHW; Bioinformatics sequence analysis: KW, TB, BK; STD NMR: JB, CB; provision of PGT122: DB, WK; provision of 35O22: M Connors. All authors read and approved the manuscript.

## **COMPETING FINANCIAL INTERESTS**

The authors declare no competing financial interests. Reprints permissions information are available at [www.cell.com](http://www.cell.com).



## REFERENCES

- Allan, J.S., Coligan, J.E., Barin, F., McLane, M.F., Sodroski, J.G., Rosen, C.A., Haseltine, W.A., Lee, T.H., and Essex, M. (1985). Major glycoprotein antigens that induce antibodies in AIDS patients are encoded by HTLV-III. *Science* 228, 109 1-1094.
- Correia, B.E., Bates, J.T., Loomis, R.J., Baneyx, G., Carrico, C., Jardine, J.G., Rupert, P., Correnti, C., Kalyuzhniy, O., Vittal, V., *et al.* (2014). Proof of principle for epitope-focused vaccine design. *Nature* 507, 201-206.
- Crooks, E.T., Tong, T., Chakrabarti, B., Narayan, K., Georgiev, I.S., Menis, S., Huang, X., Kulp, D., Osawa, K., Muranaka, J., *et al.* (2015). Vaccine-Elicited Tier 2 HIV-1 Neutralizing Antibodies Bind to Quaternary Epitopes Involving Glycan-Deficient Patches Proximal to the CD4 Binding Site. *PLoS Pathog* 11, e1004932.
- de Taeye, S.W., Ozorowski, G., Torrents de la Pena, A., Guttman, M., Julien, J.P., van den Kerkhof, T.L., Burger, J.A., Pritchard, L.K., Pugach, P., Yasmeeen, A., *et al.* (2015). Immunogenicity of Stabilized HIV-1 Envelope Trimers with Reduced Exposure of Non-neutralizing Epitopes. *Cell* 163, 1702-1715.
- Diskin, R., Klein, F., Horwitz, J.A., Halper-Stromberg, A., Sather, D.N., Marcovecchio, P.M., Lee, T., West, A.P., Jr., Gao, H., Seaman, M.S., *et al.* (2013). Restricting HIV-1 pathways for escape using rationally designed anti-HIV-1 antibodies. *J Exp Med* 210, 123 5-1249.
- Doria-Rose, N.A., Schramm, C.A., Gorman, J., Moore, P.L., Bhiman, J.N., DeKosky, B.J., Ernandes, M.J., Georgiev, I.S., Kim, H.J., Pancera, M., *et al.* (2014). Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies. *Nature* 509, 55-62.
- Dosenovic, P., von Boehmer, L., Escolano, A., Jardine, J., Freund, N.T., Gitlin, A.D., McGuire, A.T., Kulp, D.W., Oliveira, T., Scharf, L., *et al.* (2015). Immunization for HIV-1 Broadly Neutralizing Antibodies in Human Ig Knockin Mice. *Cell* 161, 1505-15 15.
- Engh R. A., H.R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallographica A* 47, 392 400.
- Garces, F., Lee, J.H., de Val, N., Torrents de la Pena, A., Kong, L., Puchades, C., Hua, Y., Stanfield, R.L., Burton, D.R., Moore, J.P., *et al.* (2015). Affinity Maturation of a Potent Family of HIV Antibodies Is Primarily Focused on Accommodating or Avoiding Glycans. *Immunity* 43, 1053-1063.
- Garces, F., Sok, D., Kong, L., McBride, R., Kim, H.J., Saye-Francisco, K.F., Julien, J.P., Hua, Y., Cupo, A., Moore, J.P., *et al.* (2014). Structural evolution of glycan recognition by a family of potent HIV antibodies. *Cell* 159, 69-79.
- 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, 8 1-90.
- Guttman, M., Garcia, N.K., Cupo, A., Matsui, T., Julien, J.P., Sanders, R.W., Wilson, I.A., Moore, J.P., and Lee, K.K. (2014). CD4-induced activation in a soluble HIV-1 Env trimer. *Structure* 22, 974-984.

- Hanover, J.A., and Lennarz, W.J. (1980). N-Linked glycoprotein assembly. Evidence that oligosaccharide attachment occurs within the lumen of the endoplasmic reticulum. *J Biol Chem* 255, 3600-3604.
- Helle, F., Duverlie, G., and Dubuisson, J. (2011). The hepatitis C virus glycan shield and evasion of the humoral immune response. *Viruses* 3, 1909-1932.
- Jardine, J., Julien, J.P., Menis, S., Ota, T., Kalyuzhniy, O., McGuire, A., Sok, D., Huang, P.S., MacPherson, S., Jones, M., *et al.* (2013). Rational HIV immunogen design to target specific germline B cell receptors. *Science* 340, 711-716.
- Jardine, J.G., Ota, T., Sok, D., Pauthner, M., Kulp, D.W., Kalyuzhniy, O., Skog, P.D., Thinnies, T.C., Bhullar, D., Briney, B., *et al.* (2015). HIV-1 VACCINES. Priming a broadly neutralizing antibody response to HIV-1 using a germline-targeting immunogen. *Science* 349, 156-161.
- Jenkins, G.M., Rambaut, A., Pybus, O.G., and Holmes, E.C. (2002). Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. *Journal of molecular evolution* 54, 156-165.
- Julien, J.P., Cupo, A., Sok, D., Stanfield, R.L., Lyumkis, D., Deller, M.C., Klasse, P.J., Burton, D.R., Sanders, R.W., Moore, J.P., *et al.* (2013). Crystal structure of a soluble cleaved HIV-1 envelope trimer. *Science* 342, 1477-1483.
- Julien, J.P., Lee, J.H., Ozorowski, G., Hua, Y., Torrents de la Pena, A., de Taeye, S.W., Nieusma, T., Cupo, A., Yasmeen, A., Golabek, M., *et al.* (2015). Design and structure of two HIV-1 clade C SOSIP.664 trimers that increase the arsenal of native-like Env immunogens. *Proc Natl Acad Sci U S A* 112, 11947-11952.
- Kong, L., Lee, J.H., Doores, K.J., Murin, C.D., Julien, J.P., McBride, R., Liu, Y., Marozsan, A., Cupo, A., Klasse, P.J., *et al.* (2013). Supersite of immune vulnerability on the glycosylated face of HIV-1 envelope glycoprotein gp120. *Nat Struct Mol Biol* 20, 796-803.
- Kong, L., Wilson, I.A., and Kwong, P.D. (2015). Crystal structure of a fully glycosylated HIV-1 gp120 core reveals a stabilizing role for the glycan at Asn262. *Proteins* 83, 590-596.
- Kornfeld, R., and Kornfeld, S. (1985). Assembly of asparagine-linked oligosaccharides. *Annu Rev Biochem* 54, 63 1-664.
- Kwon, Y.D., Pancera, M., Acharya, P., Georgiev, I.S., Crooks, E.T., Gorman, J., Joyce, M.G., Guttman, M., Ma, X., Narpala, S., *et al.* (2015). Crystal structure, conformational fixation and entry-related interactions of mature ligand-free HIV-1 Env. *Nat Struct Mol Biol* 22, 522-53 1.
- Lee, J.H., de Val, N., Lyumkis, D., and Ward, A.B. (2015). Model Building and Refinement of a Natively Glycosylated HIV-1 Env Protein by High-Resolution Cryoelectron Microscopy. *Structure* 23, 1943-1951.
- 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.
- Leonard, C.K., Spellman, M.W., Riddle, L., Harris, R.J., Thomas, J.N., and Gregory, T.J. (1990). Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells. *J Biol Chem* 265, 10373-10382.

- Lyumkis, D., Julien, J.P., de Val, N., Cupo, A., Potter, C.S., Klasse, P.J., Burton, D.R., Sanders, R.W., Moore, J.P., Carragher, B., *et al.* (2013). Cryo-EM structure of a fully glycosylated soluble cleaved HIV-1 envelope trimer. *Science* *342*, 1484-1490.
- 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.
- Mouquet, H., Scharf, L., Euler, Z., Liu, Y., Eden, C., Scheid, J.F., Halper-Stromberg, A., Gnanapragasam, P.N., Spencer, D.I., Seaman, M.S., *et al.* (2012). Complex-type N-glycan recognition by potent broadly neutralizing HIV antibodies. *Proc Natl Acad Sci U S A* *109*, E3268-3277.
- Pancera, M., Zhou, T., Druz, A., Georgiev, I.S., Soto, C., Gorman, J., Huang, J., Acharya, P., Chuang, G.Y., Ofek, G., *et al.* (2014). Structure and immune recognition of trimeric pre-fusion HIV-1 Env. *Nature* *514*, 455-461.
- Pejchal, R., Doores, K.J., Walker, L.M., Khayat, R., Huang, P.S., Wang, S.K., Stanfield, R.L., Julien, J.P., Ramos, A., Crispin, M., *et al.* (2011). A potent and broad neutralizing antibody recognizes and penetrates the HIV glycan shield. *Science* *334*, 1097-1103.
- Pritchard, L.K., Spencer, D.I., Royle, L., Bonomelli, C., Seabright, G.E., Behrens, A.J., Kulp, D.W., Menis, S., Krumm, S.A., Dunlop, D.C., *et al.* (2015). Glycan clustering stabilizes the mannose patch of HIV-1 and preserves vulnerability to broadly neutralizing antibodies. *Nature communications* *6*, 7479.
- Pugach, P., Ozorowski, G., Cupo, A., Ringe, R., Yasmeen, A., de Val, N., Derking, R., Kim, H.J., Korzun, J., Golabek, M., *et al.* (2015). A native-like SOSIP.664 trimer based on an HIV-1 subtype B env gene. *J Virol* *89*, 3380-3395.
- Sanders, R.W., Derking, R., Cupo, A., Julien, J.P., Yasmeen, A., de Val, N., Kim, H.J., Blattner, C., de la Pena, A.T., Korzun, J., *et al.* (2013). A next-generation cleaved, soluble HIV-1 Env Trimer, BG505 SOSIP.664 gp140, expresses multiple epitopes for broadly neutralizing but not non-neutralizing antibodies. *PLoS Pathog* *9*, e1003618.
- Sanders, R.W., van Gils, M.J., Derking, R., Sok, D., Ketkar, T.J., Burger, J.A., Ozorowski, G., Cupo, A., Simonich, C., Goo, L., *et al.* (2015). HIV-1 VACCINES. HIV-1 neutralizing antibodies induced by native-like envelope trimers. *Science* *349*, aac4223.
- Sanders, R.W., Venturi, M., Schiffner, L., Kalyanaraman, R., Kattinger, H., Lloyd, K.O., Kwong, P.D., and Moore, J.P. (2002). The mannose-dependent epitope for neutralizing antibody 2G12 on human immunodeficiency virus type 1 glycoprotein gp120. *J Virol* *76*, 7293-7305.
- Scanlan, C.N., Pantophlet, R., Wormald, M.R., Ollmann Saphire, E., Stanfield, R., Wilson, I.A., Kattinger, H., Dwek, R.A., Rudd, P.M., and Burton, D.R. (2002). The broadly neutralizing anti-human immunodeficiency virus type 1 antibody 2G12 recognizes a cluster of 1-2 mannose residues on the outer face of gp120. *J Virol* *76*, 7306-7321.
- Scharf, L., Wang, H., Gao, H., Chen, S., McDowall, A.W., and Bjorkman, P.J. (2015). Broadly Neutralizing Antibody 8ANC195 Recognizes Closed and Open States of HIV-1 Env. *Cell* *162*, 1379-1390.
- Sommerstein, R., Flatz, L., Remy, M.M., Malinge, P., Magistrelli, G., Fischer, N., Sahin, M., Bergthaler, A., Igonet, S., Ter Meulen, J., *et al.* (2015). Arenavirus Glycan Shield Promotes Neutralizing Antibody Evasion and Protracted Infection. *PLoS Pathog* *11*, e1005276.

- Stowell, S.R., Arthur, C.M., Mehta, P., Slanina, K.A., Blixt, O., Leffler, H., Smith, D.F., and Cummings, R.D. (2008). Galectin-1, -2, and -3 exhibit differential recognition of sialylated glycans and blood group antigens. *J Biol Chem* 283, 10109-10123.
- Szakonyi, G., Klein, M.G., Hannan, J.P., Young, K.A., Ma, R.Z., Asokan, R., Holers, V.M., and Chen, X.S. (2006). Structure of the Epstein-Barr virus major envelope glycoprotein. *Nat Struct Mol Biol* 13, 996-1001.
- Tran, E.E., Borgnia, M.J., Kuybeda, O., Schauder, D.M., Bartesaghi, A., Frank, G.A., Sapiro, G., Milne, J.L., and Subramaniam, S. (2012). Structural mechanism of trimeric HIV-1 envelope glycoprotein activation. *PLoS Pathog* 8, e1002797.
- Varki, A. (2011). Evolutionary forces shaping the Golgi glycosylation machinery: why cell surface glycans are universal to living cells. *Cold Spring Harbor perspectives in biology* 3.
- Wardemann, H., Yurasov, S., Schaefer, A., Young, J.W., Meffre, E., and Nussenzweig, M.C. (2003). Predominant autoantibody production by early human B cell precursors. *Science* 301, 1374-1377.
- Wei, X., Decker, J.M., Wang, S., Hui, H., Kappes, J.C., Wu, X., Salazar-Gonzalez, J.F., Salazar, M.G., Kilby, J.M., Saag, M.S., *et al.* (2003). Antibody neutralization and escape by HIV-1. *Nature* 422, 307-312.
- Wei, X., Ghosh, S.K., Taylor, M.E., Johnson, V.A., Emini, E.A., Deutsch, P., Lifson, J.D., Bonhoeffer, S., Nowak, M.A., Hahn, B.H., *et al.* (1995). Viral dynamics in human immunodeficiency virus type 1 infection. *Nature* 373, 117-122.
- Wyatt, R., and Sodroski, J. (1998). The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens. *Science* 280, 1884-1888.
- Zhou, T., Lynch, R.M., Chen, L., Acharya, P., Wu, X., Doria-Rose, N.A., Joyce, M.G., Lingwood, D., Soto, C., Bailer, R.T., *et al.* (2015). Structural Repertoire of HIV-1-Neutralizing Antibodies Targeting the CD4 Supersite in 14 Donors. *Cell* 161, 1280-1292.
- Zhou, T., Zhu, J., Wu, X., Moquin, S., Zhang, B., Acharya, P., Georgiev, I.S., Altae-Tran, H.R., Chuang, G.Y., Joyce, M.G., *et al.* (2013). Multidonor analysis reveals structural elements, genetic determinants, and maturation pathway for HIV-1 neutralization by VRC01-class antibodies. *Immunity* 39, 245-258.

## FIGURE LEGENDS

**Figure 1. Crystal structure of a fully glycosylated HIV-1 Env trimer at 3.4 Å resolution reveals an ordered glycan shield.**

(A) Lattice comprising Fabs 35O22 (gold) and PGT122 (light blue) holds fully glycosylated 11IV-1 Env trimers (surface representation in dark gray, with glycan highlighted in green on a representative trimer), with variable domain-only VRC01 in dark gray. (B) Crystal structure of a fully glycosylated SOSIP trimer from clade G strain X1 193.c1 with gp120 (light gray) and gp41 (dark gray) in ribbon representation and glycans (green) in stick representation. 2Fo-Fc electron density (slate blue) is shown at 0.8 for glycans; altogether over half of the carbohydrate has been crystallographically resolved. Insets show select clusters of N-linked glycans, comprising N-acetylglucosamine residues protruding perpendicularly from the protein surface and supporting mannose branches, which form a glycan canopy ~20 Å from the protein surface. A view down the 3-fold is provided in Figure 2D. See also Figure S1 and Tables S3 and S4.

**Figure 2. Fully glycosylated HIV-1 envelopes from clades A and B at 3.7 Å reveal conservation and diversity of protein and glycan shield.**

(A-B) 11IV-1 trimers with protein in ribbon representation and glycan (BG505; blue, JR-FL; magenta) in sticks. 2Fo-Fc electron density (slate blue) is shown at 0.8 for glycans; assuming the average glycan is Man-7, ~50% of the glycan mass is ordered. (C-D) Superposition of Env trimers from clades A, B and G. Protein and glycan are colored as indicated, except that the clade G protein is shown in gray for clarity. (E-F) Env conserved core (black) with structural variation

of  $C > 1.5 \text{ \AA}$  highlighted in red. **(G)** Residue-level schematic of crystallographically observed glycans. See also Figure S2 and Tables S3 and S4.

**Figure 3. Taxonomy of glycan-glycan interactions that comprise the HIV-1 glycan shield.**

**(A)** Classification based on nearest interglycan sequon distance, with histograms showing glycans from each trimer (right) and from the composite of three trimers (left). **(B)** Nearest-neighbor analysis for various viral Env glycoproteins. **(C)** Glycan residue-level details for glycan-glycan interactions. Glycans are shown in stick representation, with  $2\text{Fo}-\text{Fc}$  electron density (slate blue) at  $0.8 \text{ \AA}$ . Nearest N 2 glycan sequon neighbor distance is shown in parentheses. See also Figure S3.

**Figure 4. N-glycan crowding provides a mechanism for glycan order.**

**(A)** Number of crystallographically ordered saccharide units on a particular sequon relative to the density of neighboring N-linked glycans. Radii (R) range from  $10\text{-}55 \text{ \AA}$  and are centered on N 2 of the Asn residue in the N-linked glycan sequon. There were a total of 69 occupied sequons in the three crystal structures. All correlation values were determined using a Pearson correlation (r). **(B)** Pearson correlation values (blue) and associated  $\log(p\text{-value})$  values (black) plotted for radii ranging from  $5\text{-}60 \text{ \AA}$ . The dotted black line represents a p-value of 0.001 for reference. **(C)** The number of neighboring N-linked glycans observed in crystal structures for BG505 (clade A, left), JR-FL (clade B, middle) and x1 193.c1 (clade G, right) at a radial cutoff of  $50 \text{ \AA}$ . Blue circles denote sequon positions with at least one crystallographically observed saccharide unit. Red circles denote sequon positions with glycans in contact with an antibody. The total number

of occupied sequons for BG505, JR-FL and X1 193.c1 was 21, 23, and 25 respectively. The dotted lines represent the best partitioning of the x and y variables using Fishers Exact test (see methods for details). **(D)** Crowded and dispersed glycans mapped onto the surface of each Env trimer structure (top row). Crystallographically ordered glycans were also mapped onto the surface of each Env trimer structure (bottom row). N-linked glycans are displayed in stick representation. To enhance visualization, all neighboring residues within 10 Å were also colored accordingly. **(E)** For each glycan in the three crystallized strains, the corresponding glycan crowding was determined for each of 2994 M-group sequences by threading the BG505, JR-FL and X11 93 .c 1 structures. Blue circles indicate the median glycan neighbors, thick grey bars indicate the interquartile range and thin grey bars indicate the 95% range around the median. We found significant correlation between the number of glycan neighbors in the three strains and the median neighbors for M-group sequences. See also Figure S4.

**Figure 5. Biological characteristics of crowded and dispersed HIV-1 glycans and their associated surfaces.**

**(A)** Regions associated with crowded and dispersed glycans on the X1 193.c1 trimer are shown from side (left panel) and top (three right panels) orientations, displayed as 40 Å sections through the trimer axis. **(B)** Features of crowded, dispersed and glycan-free surfaces. Correlations of crowded and dispersed glycans with glycan processing (upper left), glycan conservation (lower left), Env sequence entropy (upper right, insertion/deletion in hypervariable regions colored black) and serologic prevalence of broadly neutralizing antibodies (lower right). See also Figure S5E.

**Figure 6. Molecular dynamics simulations reveal known broadly neutralizing antibodies that recognize the pre-fusion closed trimer need to accommodate N-linked glycan.**

(A) Epitopes for known broadly neutralizing antibodies displayed on the surface of the HIV-1 Env trimer. Epitopes are colored blue if *N*-linked glycans are known to be required for recognition and red if they are not required. (B) Glycan-antibody overlap analysis derived from MD simulations of Man-5, Man-7 and Man-9 models of the HIV-1 glycan shield. Overlap is colored blue for glycans known to be required for recognition by a particular antibody and red for glycans not known to be required for recognition. (C) Distribution graphs of required and non-required glycans. (D) Frequency of overlapping glycans versus number of overlapping glycan atoms. See also Figure S6, Table S6, and Movie S1.

**Figure 7. Broadly neutralizing antibodies recognize oligosaccharides with a broad range of affinities.**

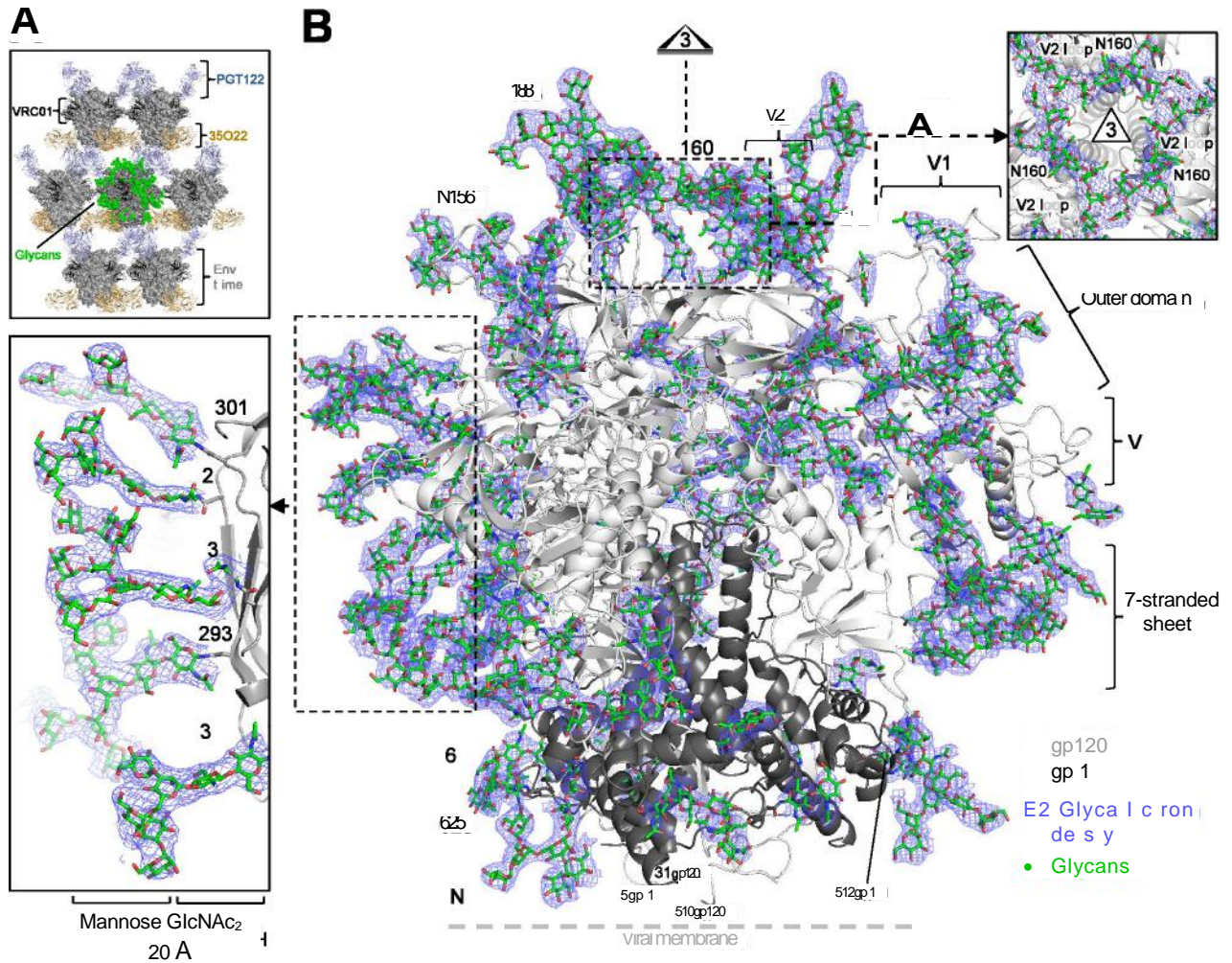
(A) Epitope-specific frequency of recognition of oligosaccharides by broadly neutralizing antibodies displayed on the surface of trimeric HIV-1 Env. (B) Relative fluorescence index ( $\times 10^6$ ) for broadly neutralizing antibody recognition of 40 oligosaccharides coupled to glass slides. (C) Ordered glycans N276 and N234 with  $2Fo-Fc$  electron density at 0.8 in the X1 193.c1

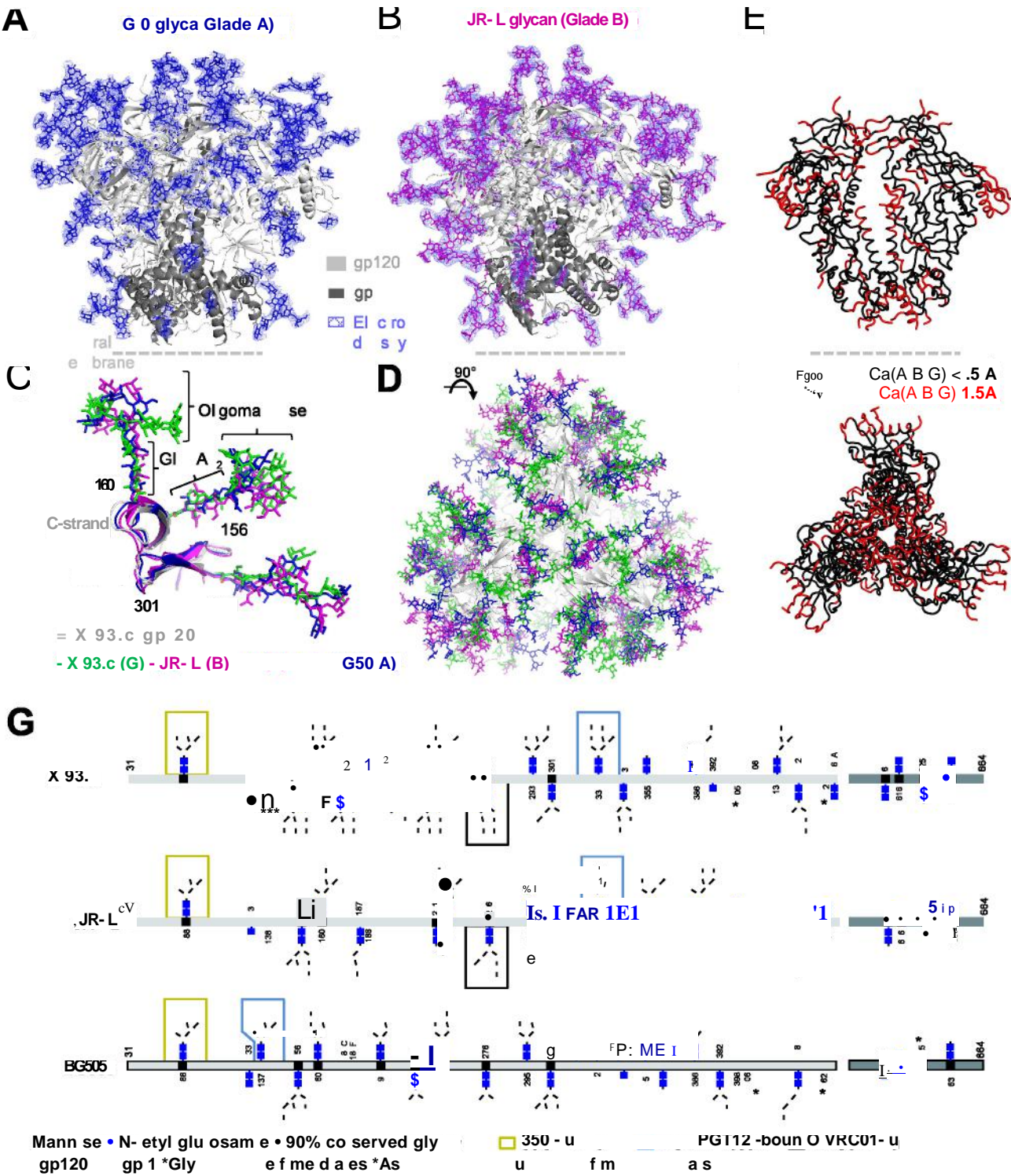
Superimposition of BG505, JR-FL and X1 193.c1 structures (middle) showing the relative difference in orientation of N276 and N234 glycans between the VRC01-bound X1 193.c1 and JR-FL structures and the BG505 structure (which was not complexed to VRC01). (D) Detection



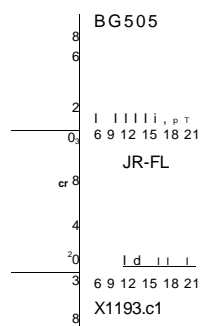
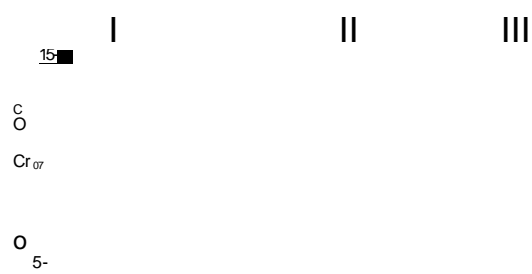
of VCR01 binding to Man<sub>9</sub>GlcNAc<sub>2</sub>Asn by Saturation Transfer Difference (STD) NMR. Signals for sugar protons are labeled above the reference spectrum (REF) (glycan only, top). STD enhancements for both N-acetyl groups of the core GlcNAc residues and pyranose protons are displayed in the difference spectrum (DIFF) (VRC01 :Man<sub>9</sub>GlcNAc<sub>2</sub>Asn complex, bottom). (E) Neutralization titers for VRC01 mutants at the N276 binding interface. (F) Molecular dynamics analysis of glycans overlapping the bound VRC01 volume, showing surrounding glycans and antibody on trimer (left) and isosurface representation of glycans overlapping VRC01 volume (right). See also Figure S7 and Table S6.

Figure 1

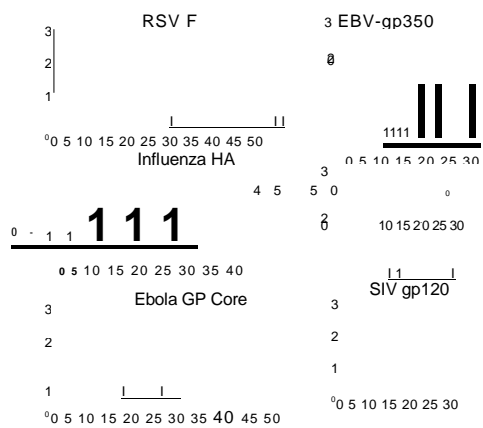




# A D e l f o



# B



0 3 , II 111111 , I , PP

4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22

D stance to nearest glycan sequon (A)

D stance to nearest glycan sequon (A)

CR u-l v l ta l

20 ||| 111

4

S -s fo k ng

-b

n

m t

Qu tern ry ax al cage

Isolated

X 93.c

3  
(A)

332  
(A)

4

S -s l g

BG505  
363  
(6 )

386  
6A

6A

Glycan l c ro de s y • Glycans

X 93.c

23  
(12 A)

1

76  
A

2

Stem-br

ch

oc at on

X 93.c

2  
(6 A)

13  
(A)

6A

3  
(A)

2

17A

120A

20A

Tr ple

gl m t

JR- L

362  
(8 A)

386  
(A)

N187  
(8 A)

6A

3A

16

Lon-r

pr ch-to-branch

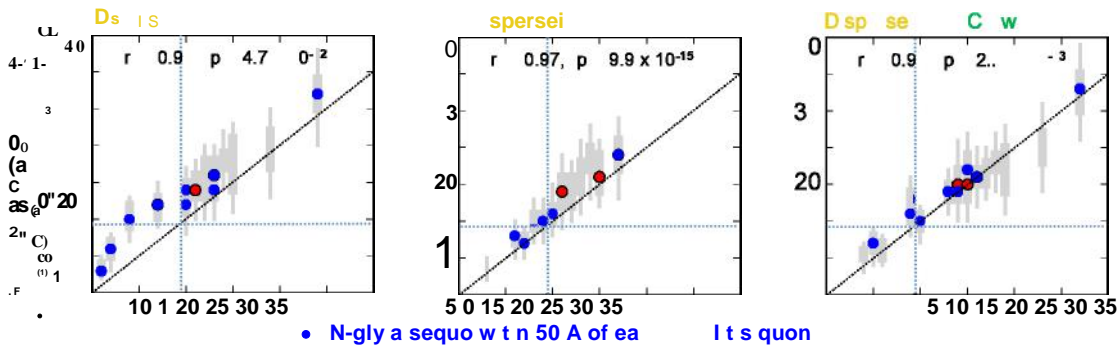
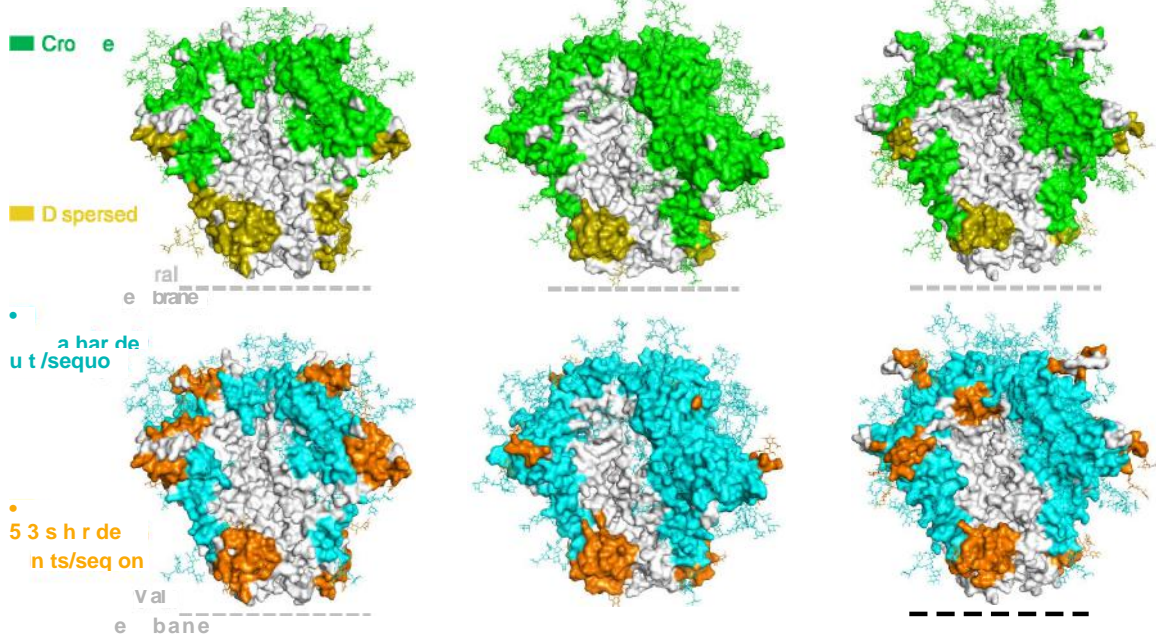
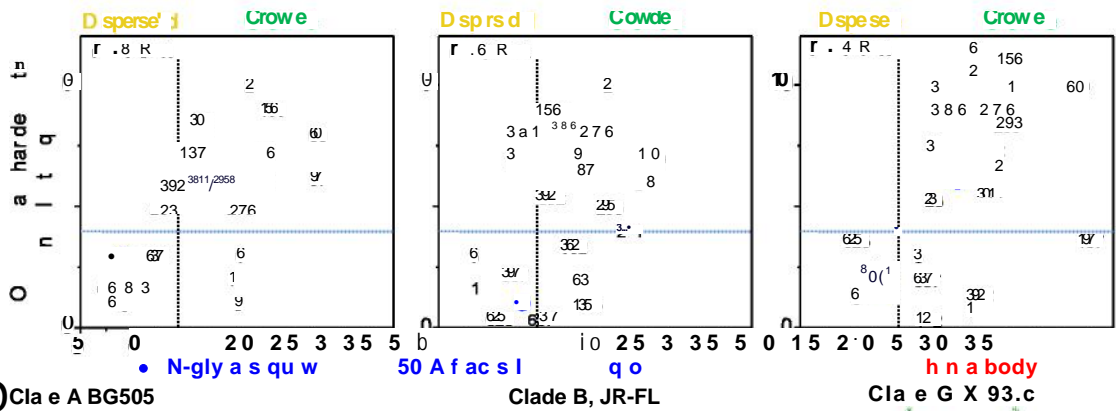
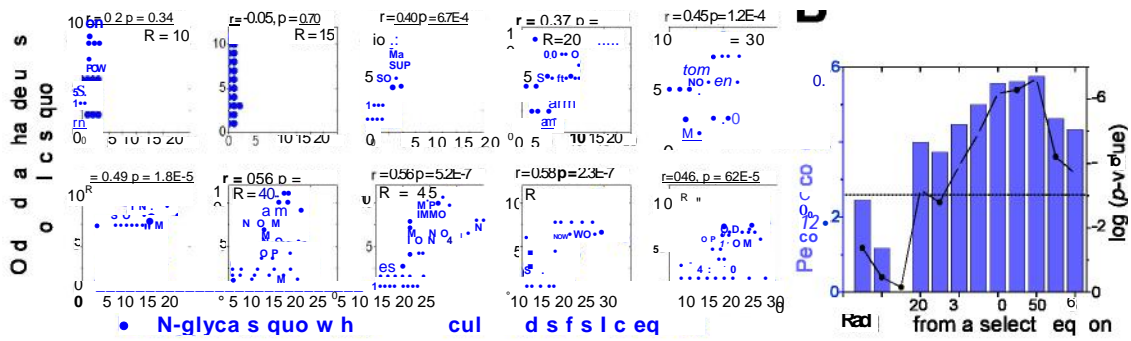
X 93.c1

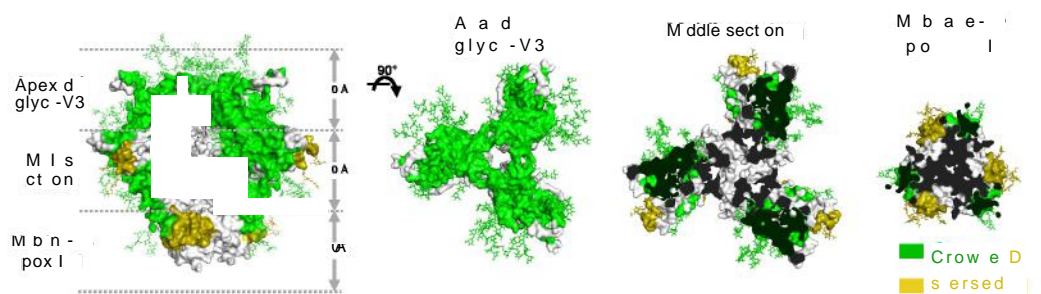
93  
(1 A)

2  
(A)

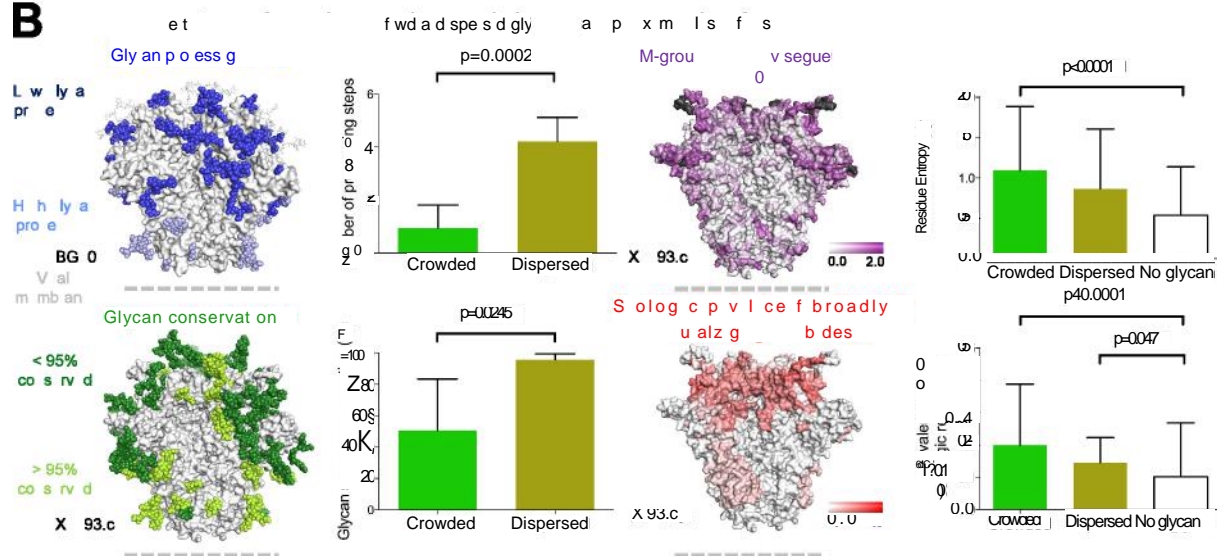
30A







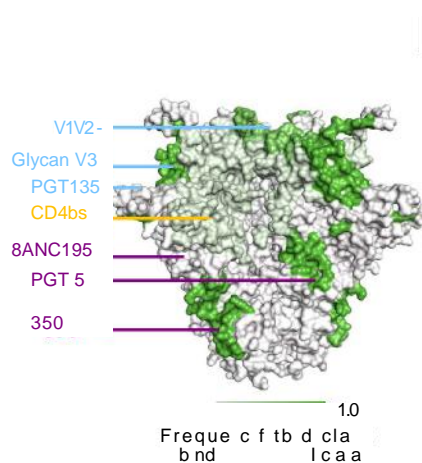
B



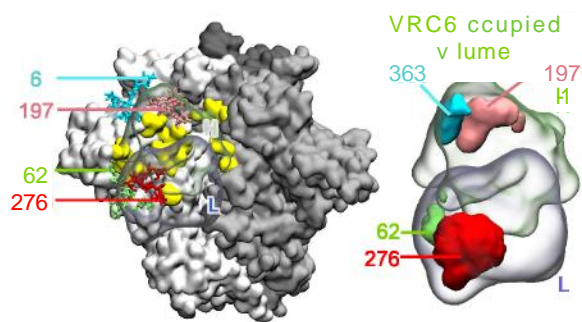
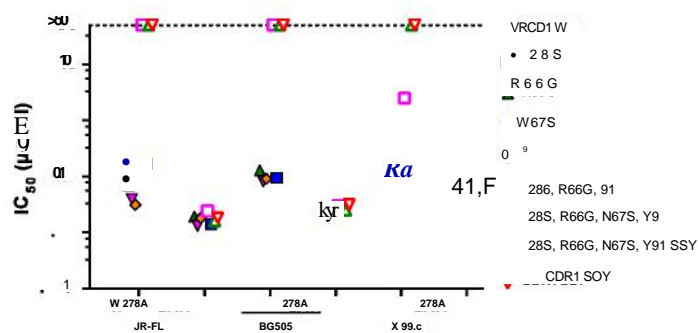
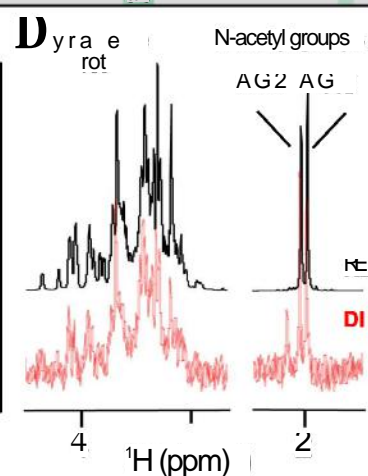
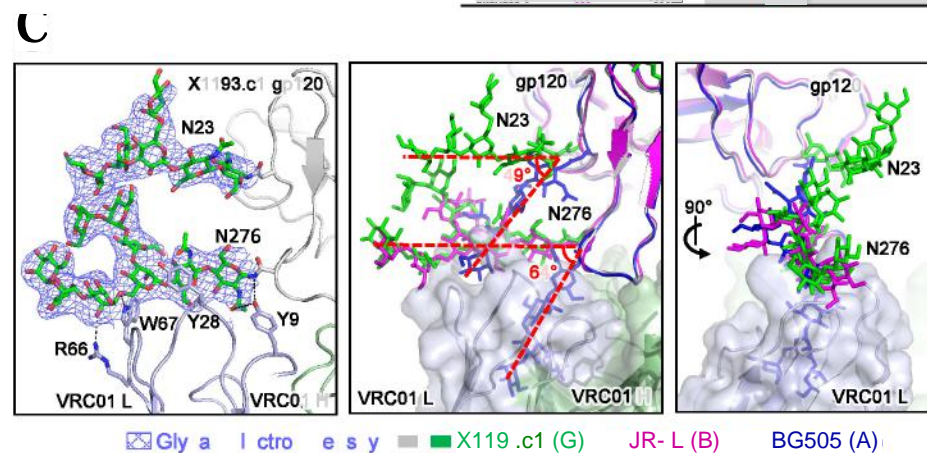




Ip O. rma ma ION a-a a  
com. Ni4e4 mmm"m m o  
z z z z 001.0gM



Bs to		Clas	Ant body	Ma ose	Hybrd	Complex		
				23	58789	1213	516789212322528222933323353833839	
12	PG9	PG9				3		
	PG5	PG5						
	PG16	PGDM100						
	R26							
Glynn-3	PG27	PG						
	29							
	PC	PG28						
3-	PG	PG36						
		b2						
		3						
		J6						
0-565ate	RC3	R3				2	3	
	RC16	R						
	RCS	RC61						
	RCO	RC-PG0						
	VRCO	RC-3						
	RCO	RC-PG20						
	RS	3050						
	RCO	22						
	RO	RC8b						
	RCO	RC23						
	RO	RC27						
	8C3	192630						
	8C3	0NC31						
	INC3	8C34						
	Interface	PG5	PG5					
		8C36	8C19b					
360		36022						



## Supplemental Information

### Trimeric HIV-1-Env Structures Define Glycan Shields from Clades A, B and G

Guillaume B. E. Stewart-Jones<sup>1</sup>, Cinque Soto<sup>1</sup>, Thomas Lemmin<sup>1,2</sup>, Gwo-Yu Chuang<sup>1</sup>, Aliaksandr Druz<sup>1</sup>, Rui Kong<sup>1</sup>, Paul V. Thomas<sup>1</sup>, Kshitij Wagh<sup>3</sup>, Tongqing Zhou<sup>1</sup>, Anna-Janina Behrens<sup>4</sup>, Tatsiana Bylund<sup>1</sup>, Chang W. Choi<sup>1</sup>, Jack R. Davison<sup>5</sup>, Ivelin S. Georgiev<sup>1</sup>, M. Gordon Joyce<sup>1</sup>, Young Do Kwon<sup>1</sup>, Marie Pancera<sup>1</sup>, Justin Taft<sup>1</sup>, Yongping Yang<sup>1</sup>, Baoshan Zhang<sup>1</sup>, Sachin S. Shivatare<sup>6</sup>, Vidya S. Shivatare<sup>6</sup>, Chang-Chun D. Lee<sup>6</sup>, Chung-Yi Wu<sup>6</sup>, Carole A. Bewley<sup>5</sup>, Dennis R. Burton<sup>7,8</sup>, Wayne C. Koff<sup>9</sup>, Mark Connors<sup>10</sup>, Max Crispin<sup>4</sup>, Ulrich Baxa<sup>11</sup>, Bette T. Korber<sup>3</sup>, Chi-Huey Wong<sup>6,12</sup>, John R. Mascola<sup>1</sup> and Peter D. Kwong<sup>1\*</sup>

<sup>1</sup> Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA.

<sup>2</sup> Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, California 94143, USA.

<sup>3</sup> Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM 87544, USA

<sup>4</sup> Oxford Glycobiology Institute, Department of Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 3QU, United Kingdom.

<sup>5</sup> Laboratory of Bioorganic Chemistry, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland, USA.

<sup>6</sup> Genomics Research Center, Academia Sinica, 128 Academia Road, Section 2, Nankang, Taipei 115, Taiwan.

<sup>7</sup> Department of Immunology and Microbial Science, International AIDS Vaccine Initiative Neutralizing Antibody Center, Center for HIV/AIDS Vaccine Immunology and Immunogen Discovery, The Scripps Research Institute, La Jolla, CA 92037, USA.

<sup>8</sup> Ragon Institute of Massachusetts General Hospital, Massachusetts Institute of Technology and Harvard University, Boston, MA 02142, USA.

<sup>9</sup> International AIDS Vaccine Initiative, New York, NY 10038, USA.

<sup>10</sup> HIV-Specific Immunity Section, Laboratory of Immunoregulation, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA.

<sup>11</sup> Electron Microscopy Laboratory, Cancer Research Technology Program, Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, Frederick, Maryland 21702, United States of America

<sup>12</sup> Department of Chemistry, The Scripps Research Institute, La Jolla, CA 92037, USA.

## EXTENDED EXPERIMENTAL PROCEDURES

### HIV-1 gp140 protein production

The production of pre-fusion closed trimers from diverse HIV-1 strains has been an obstacle to structural and immunogenicity studies. One clade A strain, BG505 (Wu et al., 2006), stabilized by a disulfide (SOS) (Binley et al., 2000), isoleucine to proline mutation (IP) (Sanders et al., 2002), and truncation at residue 664 (Julien et al., 2013), has been shown to form stable soluble trimer (Sanders et al., 2013).

Plasmids encoding HIV-1 gp140 Env designs were screened by micro-transfection in 96-well format for expression as described in (McLellan et al., 2013a). By combining plasmids for scFv VRC01 or Fab heavy and light chains from VRC01 with JR-FL or BG505 gp140 constructs, production of stable pre-fusion HIV-1 trimers was confirmed by ELISA binding with quaternary antibodies PGT145 and CAP256-VRC26.09. Mutations that allow for efficient formation of a disulfide bond at the interface of VRC01 antibody (mutations A60<sub>CHC</sub> or R61<sub>CHC</sub>) and gp140 (G459C<sub>gp120</sub>) were identified from analyzing the gp120-VRC01 complex (PDB IDs 4LSS, 4NCO, more) with the program Disulfide by Design 2.0 (Craig and Dombkowski, 2013) (**Figure S1A, B**). Co-expression of A60<sub>CHC</sub> in the antigen-binding fragment (Fab) of antibody VRC01 along with the G459C<sub>gp120</sub> mutant BG505 IP.664.T332N resulted in homogenous disulfide bond formation between antibody and trimeric Env. SDS-PAGE analysis revealed the gp120-gp41 protease site to be fully cleaved (**Figure S1C, D**). A variety of Envs from diverse HIV-1 strains were then chosen based on neutralization sensitivity to both VRC01 and PGT145, SOSIP and G459C<sub>gp120</sub> mutations were introduced, terminating at either residue 664<sub>gp41</sub> or 665<sub>gp41</sub>, and co-expressed in HEK 293 GnTI<sup>-</sup> cells with the single chain variable domains (scFv) or Fab of VRC01 with A60<sub>CHC</sub> mutation. Thirteen purified complexes

of scFv VRC01 with HIV-1 trimers from clades A, B, C, D, G, CRF\_AE and A/G/I yielded fully cleaved molecules (**Figure S1E-H**), which eluted on size-exclusion chromatography at an elution volume corresponding to ~500 kDa, and eleven of these appeared homogenous by negative-stain electron microscopy (**Figure S1I**). Antigenically, these VRC01 trimer complexes bound trimer apex quaternary antibodies PG9, PGT145, and CAP256-VRC26.09, in a manner consistent the neutralization sensitivity of the strains to these antibodies (**Figure S1J**). DNA plasmids used for co-transfection of a 1 liter cell culture (HEK 293T GnTI<sup>-</sup> or EXPI cells) were mixtures typically of 333 µg gp140, 333 µg furin, 1.67 µg VRC01 light chain and 1.67 µg VRC01 heavy chain Fab. For A60CHC scFv VRC01 co-expression, 333 µg of the plasmid was used in place of the heavy and light VRC01 antibody Fab plasmids as above. In the scFv construct the light chain was linked to the heavy chain with a (G4S)<sub>3</sub> linker between VRC01 residues 106<sub>LC</sub> and 1<sub>HC</sub>. A C-terminal, thrombin-cleavable, Ni<sup>2+</sup> resin capture purification tag (GGLVPRGSHHHHHHHH) was added after A60CHC scFv VRC01 residue 11<sub>1HC</sub>. The VRC01 Fab co-expression constructs had the same tag on the VRC01 heavy chain following position 21<sub>8HC</sub>, however for all structures reported here, the A60CHC scFv VRC01 construct was used.

For gp140 constructs, the protease cleavage site between gp120 and gp41 was mutated (HXB2 sequence REKR) to RRRRRR and either the IP mutation alone or the SOSIP mutations (Binley et al., 2000; Sanders et al., 2002) were introduced in addition to a thrombin cleavable Strep-tag II tag (GGGLVPRGGGSAWSHPQFEK) following either position 664 (including JR-FL) or 665 (including X1 193.c1) of gp140.

HEK 293T GnTI<sup>-</sup> cultures were co-transfected using TrueFect-Max transfection reagent and fed with fresh 293FreeStyle media (Life Technologies) 4 h post-transfection and with HyClone SFM4HEK293 enriched medium (HyClone) containing valproic acid (4 mM final

concentration) 24 h after transfection. Cultures were then incubated at 33° C for 6 days,

-1 Env trimer-VRC01

complexes purified by sequential  $\text{Ni}^{2+}$  (cOmplete His-Tag Purification Resin, Roche) and Strep-Tactin Superflow purification steps. Purification tags were removed by an overnight digestion with thrombin (Novagen) at 1.32 U/ml and Env trimer-VRC01 complexes were purified further by application and elution from a Superdex 200 column equilibrated with PBS.

For BG505 SOSIP.664 production, the protein was produced as described previously (Pancera et al., 2014). Briefly, BG505 SOSIP.664 DNA plasmid was co-transfected with furin plasmid DNA at a ratio of 3:2 in HEK 293T GnTI<sup>-</sup> cells using TrueFect-Max transfection reagent protocol and the cells were cultured for 6 days at 33° C. Supernatants were harvested and clarified by centrifugation and filtration through a protein was eluted with 3M  $\text{MgCl}_2$ , 10 mM Tris pH 8.0. The eluate was concentrated to less than 2 ml with a Centricon-70 concentrator and applied to a Superdex 200 column equilibrated in 5 mM HEPES, pH 7.5, 150 mM NaCl and 0.02% azide. The peak corresponding to trimeric HIV-1 Env was pooled and kept on ice until required.

### **Fab expression and purification**

PGT122 and 35O22 IgGs were expressed as previously described (Pancera et al., 2014). Heavy chain plasmids containing the HRV3C cleavage site after Lys 218 in the hinge region were co-transfected with light chain plasmids in EXPI (35O22) or HEK 293S GnTI<sup>-</sup> (PGT 122) cells using TrueFect- protocol. Cultures were fed with fresh 293FreeStyle media (Life Technologies) 4 h post-

transfection and then with HyClone SFM4HEK293 enriched medium (HyClone) containing valproic acid (4 mM final concentration) 24 h after transfection. Cultures were incubated at 33°C for 6 days, and supernatants harvested, clarified by centrifugation and filtration and passed over a protein A affinity column. After a PBS wash and low pH elution, the eluate was pH neutralized with 1M Tris pH 8.0. Fabs were obtained using HRV3C digestion overnight at room temperature and collected from flow-through from protein A column to remove Fc fragments. Fabs were then further purified through a Superdex 200 column in PBS.

### **Fab complex preparation**

PGT122 and 35O22 Fabs were each added to a solution of purified gp140 trimers in >2-fold molar excess overnight at room temperature. The complexes were then concentrated and purified over gel filtration equilibrated in 5 mM HEPES, pH 7.5, 150 mM NaCl, 0.02% azide (**Figure S1K**). Fractions were pooled, concentrated to 6-10 mg/mL and used for crystallization screening or kept on ice until further use.

### **Crystallization screening**

The HIV-1 trimer complexes were screened for crystallization using 572 conditions from Hampton, Wizard and Precipitant Synergy (Majeed et al., 2003) screens using a Cartesian Honeybee crystallization robot as described previously (McLellan et al., 2011) and a mosquito robot using 100 nl of reservoir solution and 100 nl of protein solution. For the BG505 SOSIP.664-PGT122-3 5O22 complex, a 10% v/v additive screen (Hampton Research) was performed based on the previously reported condition for crystallizing the partially deglycosylated ternary complex: 16% isopropanol, 5.32% PEG 1500, 0.2 M Li<sub>2</sub>SO<sub>4</sub>, 0.1M Na

acetate pH 5.5. Crystals grew optimally as rods in 0.18 M Li<sub>2</sub>SO<sub>4</sub>, 4.79% PEG 1500, 14.4% isopropanol, 10 mM yttrium chloride and 0.09 M sodium acetate pH 5.5. Crystals of fully glycosylated 293T-expressed BG505 SOSIP bound to PGT122 and 35O22 Fabs were also obtained; however the quality of these crystals needed further optimization to be sufficient for atomic-level analysis. For both the JR-FL-scFv VRC01-PGT122-35O22 and X1 193.c1-scFv VRC01-PGT122-35O22, crystals were identified in the 100 nl screen with 4.95% isopropanol, 8.25% PEG 3350 and 0.2M ammonium citrate pH 4.5 and hexagonal rods grew to maximum

We were also able to obtain crystals containing clade C trimers using the PGT 122 and 35O22 lattice, but the crystal size and resolution obtained were of lower quality than with the BG505, JR-FL and X1 193.c1 complexes. Crystals were cryoprotected in solutions of 20% ethylene glycol, maintaining the mother liquor components at the original concentrations, and flash-frozen in liquid nitrogen. Data were collected at a wavelength of 1.00 Å at the SER-CAT beamline ID-22 (Advanced Photon Source, Argonne National Laboratory).

### **X-ray data collection, structure solution and model building**

Crystals for X1 193 .c1-scFv VRC01-PGT122-35O22 were spacegroup P6<sub>3</sub>, and diffraction extended to 3.1 Å resolution along the *c* axis (*c*-axis > 3.0) and to lower resolutions along *a* and *b* axes, resulting in an overall dataset resolution with of 2.0 of 3.4 Å (**Table S1**). Structure solution by molecular replacement followed by rigid body fitting revealed extensive glycosylation; rebuilding to optimize geometry and glycan stereochemistry yielded *R*<sub>work</sub>/*R*<sub>free</sub> values of 21.4%/27.2% (**Table S1, Figure S1L-O**). Diffraction from JR-FL-scFv VRC01-PGT122-35O22 and BG505-PGT122-35O22 crystals were similarly anisotropic to



the X1 193.c1 data and extended to nominal resolutions of 3.7 Å. Molecular replacement and refinement led to  $R_{work}/R_{free}$  of 25.3%/30.9% and 26.0%/30.7% for JR-FL and BG505, respectively (**Tables S1 and S3**). Interestingly, despite general similarity of the lattice, the crystallizing antibodies did show strain-dependent variation, especially with N137, a glycan known to be important in PGT122 recognition (Garces et al., 2015; Garces et al., 2014). N137 was disordered in JR-FL, ordered in BG505 and displaced by a V1 extension in X11 91 .c 1 (**Figure S2E**).

Diffraction data were processed with the HKL2000 suite (Otwinowski and Minor, 1997). The data were corrected for anisotropy by <http://services.mbi.ucla.edu/anisotropy/> with truncations to 3.4 Å, 3.4 Å, 3.1 Å (X1 193.c1) and 3.7 Å, 3.7 Å, 3.1 Å (JR-FL and BG505) along a, b, and c axes, respectively. Structure solution was obtained with Phaser (McCoy et al., 2007) using BG505 SOSIP.664-PGT122-35O22 (PDB ID: 4TVP) (Pancera et al., 2014), as search model. Protein polypeptide chains for gp120 and gp41 were re-built using Autobuilder (Terwilliger et al., 2008) to eliminate any glycan model bias and in the absence of glycan coordinates revealed extensive and unambiguous carbohydrate electron density near the protein surface from *Fo-Fc* difference maps. Refinement was carried out with Phenix (Adams et al., 2004) and Buster (Bricogne, 2011). Model building was carried out with Coot (Emsley and Cowtan, 2004). Carbohydrates were modelled into *Fo-Fc* maps and further refined using *Fo-Fc*, *2Fo-Fc* and feature-enhanced maps (Afonine et al., 2015). Glycan geometries and real-space correlation coefficients were determined using the program Privateer, confirming positive real-space correlation coefficients for all modelled glycans and 100%  $^4C_1$  carbohydrate geometry for all sugar residues in the three structures (Agirre et al., 2015). The crystallographic refinement of glycans enabled dominant glycan conformations and estimations of glycan type based on

density-fitting; however, the B-factors indicate that flexibility inherent to HIV-1 Env glycans might result in modelled structures that could vary from those present here. The Ramachandran plot as determined by MOLPROBITY (Davis et al., 2004) showed >89.8% of all residues in favored regions and >99.6% of all residues in allowed regions. Data collection and refinement statistics are shown in Table S1. Surface area calculations were performed using the program ASC (Eisenhaber and Argos, 1993).

### **Glycan analysis**

Glycosylation profiles of N-linked glycans of purified Env proteins and antibodies were determined by hydrophilic interaction liquid chromatography-ultra performance liquid chromatography (HILIC-UPLC) as previously described (Pritchard et al., 2015). Briefly, glycoprotein bands were excised from Coomassie blue stained SDS-gels and released by addition of protein N-glycosidase F (PNGase F) at 5 U/ml and incubation at 37 °C for 6 h, according to the manufacturer's instructions (New England Biolabs). Released glycans were eluted from the gel bands by washing with water, dried down using a SpeedVac concentrator, labelled with 2-aminobenzoic acid (2-AA) and purified. Fluorescently labelled glycans were resolved by HILIC-UPLC using a 2.1 mm × 10 mm Acquity BEH Amide column as previously described. Empower 3 software was used for peak integration.

### **Saturation-Transfer Difference Nuclear Magnetic Resonance**

Binding of VRC01 to Man<sub>9</sub>GlcNAc<sub>2</sub>Asn was detected by NMR using saturation transfer methods as described (McLellan et al., 2011; Meyer and Peters, 2003). In this method, ligand protons that were within 5 Å of the protein or receptor in the bound state showed STD

enhancement. Data were acquired on a Bruker Avance 600 MHz instrument equipped with a triple resonance cryo-probe with z-axis gradients. Topspin 2.1 software was used to process and analyze the NMR data. The saturation pulse consisted of a train of 50 msec EBURB2 pulses (BW=100 Hz, B 1=82 Hz) separated by 1 msec for a total saturation time of 2 sec with on and off resonance pulses applied at -1.5 and -20 ppm, respectively. A 10 KHz spin lock was applied for 10 msec before 39-19 water suppression and 1s of data acquisition (sw = 15ppm, 16k points, 1 and 2 msec post-acquisition purge pulses at 17 kHz, relaxation delay = 2.5s, 4 K scans). The saturation transfer difference spectrum was obtained by subtracting the on-resonance spectrum from the reference spectrum. Samples were prepared in 20 mM sodium phosphate buffer containing 50 mM sodium chloride at pH 6.8. All experiments were carried out at 298 K.

### **Nearest Inter-N-linked glycan sequon analysis**

The distances for each N-linked glycan sequon and nearest N-linked glycan sequon were calculated for various viral glycoproteins. Asparagine N 2 atoms were used to calculate the distances. We analyzed respiratory syncytial virus (RSV) fusion glycoprotein (McLellan et al., 2013b), influenza virus hemagglutinin (HA) (Stevens et al., 2006) and the Ebola virus core envelope glycoprotein (GP) (Lee et al., 2008), where glycan sequons were considerably more sparse than with HIV-1 (**Figure 3B**). Higher sequon density was observed with SIV gp120 (Chen et al., 2005), HCV E2 core (Kong et al., 2015), and the Epstein-Barr virus gp350 (Szakonyi et al., 2006), though none of these displayed glycans with nearest neighbor distances of less than 9 Å. The PDB IDs used for the non-HIV-1 Env glycoproteins were the following: RSV fusion protein: 4JHW; influenza HA: 4M4Y; Ebola GP core: 3CSY; EBV-gp350: 2H6O; HCV E2 core: 4MWF; SIV gp120: 2BF1.

## Glycan modeling and molecular dynamics simulations

We constructed three all-atom fully glycosylated trimer models using the coordinates of the BG505 x-ray structure (PDB ID: 5FYL) as a template. Any missing segments in the BG505 x-ray structure were modeled using LoopyMod (Soto et al., 2008) based on the native BG505 sequence. Missing side chains were modeled using the program SCAP (Xiang and Honig, 2001). Using in-house software we modeled a mannose-5, mannose-7 and mannose-9 (Man-5, Man-7 and Man-9) sugar moiety at each N-linked glycosylation sequon along the BG505 protein sequence. The resulting three models (Man-5, Man-7 and Man-9) were fully solvated in a water box of dimensions  $175 \times 175 \times 175 \text{ \AA}^3$  with a  $17 \text{ \AA}$  padding. The system was neutralized by the addition of NaCl at a concentration of 150 mM. The CHARMM36 force field was used for the parameterization of the protein (including CMAP corrections) (Best et al., 2012) and the mannose-9 (Guvench et al., 2011). TIP3P parameterization (Jorgensen, 1983) was used to describe the water molecules.

All molecular dynamics (MD) simulations involving the three fully glycosylated trimer models (Man-5, Man-7 and Man-9) were performed with ACEMD software (Harvey et al., 2009) in explicit solvent on a METROCUBO workstation (<https://www.acellera.com/products/GPU-hardware-molecular-dynamics-metrocubo/>). The system was minimized for 2000 steps, followed by equilibration in the NPT ensemble for 50 ns at 1 atm and 300 K using a time-step of 2 fs, rigid bonds, cutoff of  $9 \text{ \AA}$  and PME for long range electrostatics. During the equilibration, heavy protein atoms were constrained by a  $1 \text{ kcal/mol\AA}^2$  spring constant and slowly relaxed over the first 5 ns, and the protein was allowed to move freely thereafter. The system was then simulated for 500 ns under the NVT ensemble using ACEMD in the NVT ensemble using a Langevin

thermostat with damping of  $0.1 \text{ ps}^{-1}$  and hydrogen mass repartitioning scheme to achieve time-steps of 4 fs.

Snapshots from each of the three the MD simulation were taken at 0.1 ns time intervals for structural analysis. To assess the internal geometry of each glycan moiety, we analyzed the distribution of GlcNAc-GlcNAc- -mannose angles (kink angle) for each simulation. The kink angle is defined as the angle formed between the geometric centers of the first three saccharide  
ated asparagine residue. We also assessed the distribution of glycan-glycan contacts by a defining a contact as any glycan that has 10 or more atoms within a distance of  $5.0 \text{ \AA}$  to another glycan.

### **Glycan-antibody overlap analysis**

For the glycan-antibody overlap analysis we determined the number of atoms from each glycan that occupied the same volume that would be potentially occupied by an antibody. We used 8 structures that were co-crystallized with gp120 (PDB IDs: 4JAN, 4JM2, 4YDJ, 4YE4, 3U2S, 5FYJ, 4P9H and 2NY7), two antibody structures that were co-crystallized an HIV-1 prefusion trimer (PGT122 and 35O22; PDB ID: 4TVP), one antibody structure docked onto a HIV-1 prefusion trimer using a cryoEM density map (PGT151; PDB ID: 4NUG) and one antibody structure co-crystallized with CD4 (PDB ID: 1GC1). Each co-crystallized structure was aligned to each of the three MD trajectories at 0.1 ns time points using TM-align (Zhang and Skolnick, 2005). After superimposition, we determined all glycan atoms from the MD trajectory within  $3.0 \text{ \AA}$  of the antibody or CD4 structures.

## **Antibody epitope definition and serological prevalence**

The antibody-epitope residues are antigen residues that have at least one non-hydrogen atom within 5.5 Å of any antibody non-hydrogen atom. The chronic serological prevalence of 11IV-1 nAbs based on analysis of serum neutralization of a panel of diverse 11IV-1 isolates were adapted from (Pancera et al., 2014), with detailed method described in (Georgiev et al., 2013).

## **Structural analysis and the definition of crowded and dispersed glycans**

N-linked glycosylation sequons in each of the x-ray structures that contained an attached oligosaccharide unit were labeled as occupied and used to determine the number of nearby sequons within an R angstrom spherical shell (where R ranged in length from 10-60 Å spaced at 5 Å intervals). The distance between occupied N-linked glycosylation sequons was measured occupied sequons at each position along the three protein sequences was correlated with the total number of resolved oligosaccharide units for each spherical shell using a Pearson correlation coefficient (with a two-tailed test to determine the statistical significance). All statistical calculations were carried out using R (Core, 2013).

The number of neighboring glycans from each of the three x-ray structures was also used to determine if a glycan was either crowded or dispersed. We used a  $2 \times 2$  contingency table where we defined groups as either crowded or dispersed depending on the number of neighboring sequons within a spherical shell of 50 Å and the number of attached saccharide -30 Å representing the spherical shell containing the number of neighboring sequons and 3-1 1 for the number of saccharide units attached to the sequon. We achieved optimal partitioning at a value of 15 Å and

4 saccharide units (according the computed p-value). Thus, any glycan with 15 or more neighboring sequons within a 50 Å shell was labeled as being crowded and any glycan within test was calculated using R.

To assess the variability of M-group sequences on each of the three x-ray structures we threaded 2994 M-group Env sequences from all major clades (without recombinants) from the Los Alamos HIV database ([www.hiv.lanl.gov](http://www.hiv.lanl.gov)) onto each structure using the homology modeling package Nest with default settings (Petrey et al., 2003). calculated the number of glycan sequons within a 50 Å radius of each glycan sequon to compare these values to those found for BG505, JR-FL and X1 193.c1 structures (Figure 4E, Figure S4E). The M-group crowded glycan proximal, dispersed glycan proximal, and no glycan residues were defined as follows. Only those N-glycan sequence motifs for which Asn was at a surface-exposed position ( $SASA > 0.25$ ) in the native crystal structure used for threading were considered as active sequons. The surfaces for each of the M-group sequences using the structures threaded to the X11 93 .c 1 crystal structure, and then each residue was classified according to the most prevalent type observed over all M-group sequences. Significant correlation was observed between the number of glycan neighbors observed for each crystallized strain and the median number of glycan neighbors for the M-group models (Pearson correlation coefficient  $r = 0.96$   $0.97$ ,  $p = 9.9 \times 10^{-15}$   $4.7 \times 10^{-12}$  ). Pearson correlation coefficients and associated p-values were calculated using the package Scientific Python (SciPy).

The crowded-glycan-proximal surface were defined as surface residues (residues with at least 25% surface accessibility) which have at least one heavy atom within 10 angstroms of any crow

The dispersed-glycan-proximal surface were defined as

surface residues which have at least one heavy atom within 10 Å of any dispersed glycan  
-glycan proximal surface residue.

### **Fabrication of NHS-Glycan Microarray**

The 40 glycans used in the array analysis are prepared by the modular synthesis method developed recently (Shivatara et al., 2013; Shivatara, 2015). All monovalent glycans were prepared in 10 mM concentration individually and served as mother solutions, which were diluted with printing buffer to prepare a working solution. Microarrays were prepared by printing (BioDot Cartesion Technologies) (Blixt et al., 2004) with robotic pin (SMP3: TeleChem International) and printed onto NHS-coated glass slides (Nexterion H, SCHOTT). Printed slides were allowed to react in an atmosphere of 80% humidity for one hour followed by desiccation overnight. Individual glycans were spotted with five replicas. The prepared slides were stored in a humidity controlled dry box before the binding assay. The HIV-1 broadly neutralizing antibodies were complexed with Donkey Anti-Human IgG added to the glycan array and incubated at 4°C overnight. Slides were scanned with a GenePix 4300A reader (Molecular Devices, USA) and analyzed with GenePix Pro 7.0 software (Molecular Devices, USA).

### **Antibody-Glycan Binding Assay**

Glycan arrays were blocked with blocking buffer (Superblock, Thermo Fisher Scientific, USA) for 1 hour and washed with PBST buffer (PBS buffer with 0.05% Tween 20). The HIV-1 broadly neutralizing antibodies /mL in BSA contained PBST buffer, 3% w/v) were pre-complexed with Donkey Anti-Human IgG (Alexa Fluor647 conjugated, Jackson ImmunoResearch, USA) in 1:1 ratio and incubated at 4°C for 30 minutes. The pre-complexed mixture was added to the glycan array and incubated at 4°C overnight. The slides were washed sequentially in PBST buffer (PBS and 0.05% Tween-20), and de-ionized water. Slides were spin-dried prior to scanning with a GenePix 4300A reader (Molecular Devices, USA) and analyzed



with GenePix Pro 7.0 software (Molecular Devices, USA). The image size was 5  $\mu\text{m}$  by 5  $\mu\text{m}$ . The pixel size was defined as 0.1  $\mu\text{m}$ . The total intensity of fluorescence was calculated and illustrated with Prism 6.0 (Graphpad, USA). Error bars represent the average percentage error for all data points reported (**Figure S6F, G**).

### Neutralization assays

Monoclonal antibody (mAb) neutralization was assessed in TZM-bl cells as described previously (Montefiori, 2009; Wu et al., 2010). Briefly, 293T cells were cotransfected by a and a wildtype BG505, JR-FL and X1 193.c1 or T278A mutant HIV-1 Env expression plasmids to produce Env-pseudotyped virus stocks. Viruses were mixed with 5-fold serially diluted mAbs starting at , and incubated at 37 °C for 1 hour before adding to the cells. After incubation at 37 °C for 48 hours, the supernatants were removed and the cells were lysed. Luciferase activity was measured. 50% inhibitory concentrations (IC<sub>50</sub>) were determined as described (Wu et al., 2010).

### SUPPLEMENTAL ACKNOWLEDGEMENTS

We thank Acellera for assistance with molecular dynamics simulations. made possible by support from many donors including: the Bill & Melinda Gates Foundation; the Ministry of Foreign Affairs of Denmark; Irish Aid; the Ministry of Finance of Japan; the Ministry of Foreign Affairs of the Netherlands; the Norwegian Agency for Development Cooperation (NORAD); the UK Department for International Development (DFID); and the United States Agency for International Development (USAID). The full list of IAVI donors is available at <http://www.iavi.org>. A-J. B. and M.Crispin were supported by the

International

AIDS Vaccine Initiative Neutralizing Antibody Center CAVD grant (Glycan characterization and Outer Domain glycoform design) and the Scripps Center For HIV/AIDS Vaccine Immunology-Immunogen Discovery (CHAVI-ID) grant 1UM1AI100663. K. W. and B. K. were supported by the Duke CHAVI-ID grant UM1 AI100645. T. L. was supported by the Swiss National Foundation of Science Fellowship 148914. V. S. S., C. C. L., C. Y. W. and C. H. W. were supported by Academia Sinica and Ministry of Science and Technology grants MOST 104-0210-01-09-02 and 103-2321-B-001-004 and C. H. W. was supported by NIAID grant # 5 R01 AI072 155. The Frederick National Laboratory for Cancer Research, National Institutes of Health, is supported under contract HHSN261200800001E. Use of sector 22 (Southeast Region Collaborative Access team) at the Advanced Photon Source was supported by the US Department of Energy, Basic Energy Sciences, Office of Science, under contract number W-3 1-109-Eng-38. The contents of this manuscript are the responsibility of the authors and do not necessarily reflect the views of USAID or the US Government.

## SUPPLEMENTAL REFERENCES

Adams, P.D., Gopal, K., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Pai, R.K., Read, R.J., Romo, T.D., *et al.* (2004). Recent developments in the PHENIX software for automated crystallographic structure determination. *J Synchrotron Radiat* *11*, 53-55.

Afonine, P.V., Moriarty, N.W., Mustyakimov, M., Sobolev, O.V., Terwilliger, T.C., Turk, D., Urzhumtsev, A., and Adams, P.D. (2015). FEM: feature-enhanced map. *Acta Crystallogr D Biol Crystallogr* *71*, 646-666.

Agirre, J., Davies, G., Wilson, K., and Cowtan, K. (2015). Carbohydrate anomalies in the PDB. *Nature chemical biology* *11*, 303.

Best, R.B., Zhu, X., Shim, J., Lopes, P.E., Mittal, J., Feig, M., and Mackerell, A.D., Jr. (2012). Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone phi, psi and side-chain chi(1) and chi(2) dihedral angles. *Journal of chemical theory and computation* *8*, 3257-3273.

Binley, J.M., Sanders, R.W., Clas, B., Schuelke, N., Master, A., Guo, Y., Kajumo, F., Anselma, D.J., Maddon, P.J., Olson, W.C., *et al.* (2000). A recombinant human immunodeficiency virus type 1 envelope glycoprotein complex stabilized by an intermolecular disulfide bond between the gp120 and gp41 subunits is an antigenic mimic of the trimeric virion-associated structure. *J Virol* *74*, 627-643.

Blixt, O., Head, S., Mondala, T., Scanlan, C., Huflejt, M.E., Alvarez, R., Bryan, M.C., Fazio, F., Calarese, D., Stevens, J., *et al.* (2004). Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc Natl Acad Sci U S A* *101*, 17033-17038.

Bricogne, G.B., E.; Brandl, M.; Flensburg, C.; Keller, P.; Paciorek, W.; Roversi, P.; Sharff, A.; Smart, OS.; Vonnrhein, C.; Womack, T. O. (2011). BUSTER version 2.10.0. . Cambridge, UK: Global Phasing Ltd.

Chen, B., Vogan, E.M., Gong, H., Skehel, J.J., Wiley, D.C., and Harrison, S.C. (2005). Structure of an unliganded simian immunodeficiency virus gp120 core. *Nature* *433*, 834-841.

Core, T.R. (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Craig, D.B., and Dombkowski, A.A. (2013). Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins. *BMC Bioinformatics* *14*, 346.

Davis, I.W., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2004). MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res* *32*, W615-619.

- Eisenhaber, F., and Argos, P. (1993). Improved strategy in analytical surface calculation for molecular system- handling of singularities and computational efficiency. *J Comp Chem* *14*, 1272-1280.
- Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* *60*, 2126-2132.
- Garces, F., Lee, J.H., de Val, N., Torrents de la Pena, A., Kong, L., Puchades, C., Hua, Y., Stanfield, R.L., Burton, D.R., Moore, J.P., *et al.* (2015). Affinity Maturation of a Potent Family of HIV Antibodies Is Primarily Focused on Accommodating or Avoiding Glycans. *Immunity* *43*, 1053-1063.
- Garces, F., Sok, D., Kong, L., McBride, R., Kim, H.J., Saye-Francisco, K.F., Julien, J.P., Hua, Y., Cupo, A., Moore, J.P., *et al.* (2014). Structural evolution of glycan recognition by a family of potent HIV antibodies. *Cell* *159*, 69-79.
- Georgiev, I.S., Doria-Rose, N.A., Zhou, T., Kwon, Y.D., Staup, 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*, 751756.
- Guvench, O., Mallajosyula, S.S., Raman, E.P., Hatcher, E., Vanommeslaeghe, K., Foster, T.J., Jamison, F.W., 2nd, and Mackerell, A.D., Jr. (2011). CHARMM additive all-atom force field for carbohydrate derivatives and its utility in polysaccharide and carbohydrate-protein modeling. *Journal of chemical theory and computation* *7*, 3162-3180.
- Harvey, M.J., Giupponi, G., and Fabritiis, G.D. (2009). ACEMD: Accelerating Biomolecular Dynamics in the Microsecond Time Scale. *Journal of chemical theory and computation* *5*, 1632-1639.
- Jorgensen, W.C., J.; Madura, J.; Impey, R.; Klein, M. (1983). Comparison of simple potential functions for simulating liquid water. *J Chem Phys* *79*, 926-93 5.
- 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 U S A* *110*, 435 1-4356.
- Kong, L., Kadam, R.U., Giang, E., Ruwona, T.B., Nieuwsma, T., Culhane, J.C., Stanfield, R.L., Dawson, P.E., Wilson, I.A., and Law, M. (2015). Structure of Hepatitis C Virus Envelope Glycoprotein E1 Antigenic Site 314-324 in Complex with Antibody IGH526. *J Mol Biol* *427*, 26 17-2628.
- Lee, J.E., Fusco, M.L., Hessel, A.J., Oswald, W.B., Burton, D.R., and Saphire, E.O. (2008). Structure of the Ebola virus glycoprotein bound to an antibody from a human survivor. *Nature* *454*, 177-182.
- Majeed, S., Ofek, G., Belachew, A., Huang, C.C., Zhou, T., and Kwong, P.D. (2003). Enhancing protein crystallization through precipitant synergy. *Structure* *11*, 1061-1070.
- McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. *Journal of applied crystallography* *40*, 65 8-674.

McLellan, J.S., Chen, M., Joyce, M.G., Sastry, M., Stewart-Jones, G.B., Yang, Y., Zhang, B., Chen, L., Srivatsan, S., Zheng, A., *et al.* (2013a). Structure-based design of a fusion glycoprotein vaccine for respiratory syncytial virus. *Science* *342*, 592-598.

McLellan, J.S., Chen, M., Leung, S., Graepel, K.W., Du, X., Yang, Y., Zhou, T., Baxa, U., Yasuda, E., Beaumont, T., *et al.* (2013b). Structure of RSV fusion glycoprotein trimer bound to a prefusion-specific neutralizing antibody. *Science* *340*, 1113-1117.

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.

Meyer, B., and Peters, T. (2003). NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. *Angew Chem Int Ed Engl* *42*, 864-890.

Montefiori, D.C. (2009). Measuring HIV neutralization in a luciferase reporter gene assay. *Methods Mol Biol* *485*, 395-405.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* *276*, 307-326.

Pancera, M., Zhou, T., Druz, A., Georgiev, I.S., Soto, C., Gorman, J., Huang, J., Acharya, P., Chuang, G.Y., Ofek, G., *et al.* (2014). Structure and immune recognition of trimeric pre-fusion HIV-1 Env. *Nature* *514*, 455-461.

Petrey, D., Xiang, Z., Tang, C.L., Xie, L., Gimpelev, M., Mitros, T., Soto, C.S., Goldsmith-Fischman, S., Kernytsky, A., Schlessinger, A., *et al.* (2003). Using multiple structure alignments, fast model building, and energetic analysis in fold recognition and homology modeling. *Proteins* *53 Suppl 6*, 430-435.

Pritchard, L.K., Vasiljevic, S., Ozorowski, G., Seabright, G.E., Cupo, A., Ringe, R., Kim, H.J., Sanders, R.W., Doores, K.J., Burton, D.R., *et al.* (2015). Structural Constraints Determine the Glycosylation of HIV-1 Envelope Trimers. *Cell reports* *11*, 1604-1613.

Sanders, R.W., Derking, R., Cupo, A., Julien, J.P., Yasmeen, A., de Val, N., Kim, H.J., Blattner, C., de la Pena, A.T., Korzun, J., *et al.* (2013). A next-generation cleaved, soluble HIV-1 Env Trimer, BG505 SOSIP.664 gp140, expresses multiple epitopes for broadly neutralizing but not non-neutralizing antibodies. *PLoS Pathog* *9*, e1003618.

Sanders, R.W., Vesanen, M., Schuelke, N., Master, A., Schiffner, L., Kalyanaraman, R., Paluch, M., Berkhout, B., Maddon, P.J., Olson, W.C., *et al.* (2002). Stabilization of the soluble, cleaved, trimeric form of the envelope glycoprotein complex of human immunodeficiency virus type 1. *J Virol* *76*, 8875-8889.

Shivatare, S.S., Chang, S.H., Tsai, T.I., Ren, C.T., Chuang, H.Y., Hsu, L., Lin, C.W., Li, S.T., Wu, C.Y., and Wong, C.H. (2013). Efficient convergent synthesis of bi-, tri-, and tetra-antennary complex type N-glycans and their HIV-1 antigenicity. *J Am Chem Soc* *135*, 15382-15391.

Shivatare, S.S.C., S.-H.; Tsai, T.-I.; Tseng, S. Y.; Lin, Y.-H.; Cheng, Y.-Y.; Ren, C.-T.; Lee, C.-C.; Pawar, S.; Tsai, C.-S.; (2015). Modular synthesis of N-glycans and arrays for the hetero-ligand binding analysis of HIV antibodies. *Nat Chem*, in press.

Soto, C.S., Fasnacht, M., Zhu, J., Forrest, L., and Honig, B. (2008). Loop modeling: Sampling, filtering, and scoring. *Proteins* *70*, 834-843.

- Stevens, J., Blixt, O., Tumpey, T.M., Taubenberger, J.K., Paulson, J.C., and Wilson, I.A. (2006). Structure and receptor specificity of the hemagglutinin from an H5N1 influenza virus. *Science* 312, 404-410.
- Szakonyi, G., Klein, M.G., Hannan, J.P., Young, K.A., Ma, R.Z., Asokan, R., Holers, V.M., and Chen, X.S. (2006). Structure of the Epstein-Barr virus major envelope glycoprotein. *Nat Struct Mol Biol* 13, 996-1001.
- Terwilliger, T.C., Grosse-Kunstleve, R.W., Afonine, P.V., Moriarty, N.W., Zwart, P.H., Hung, L.W., Read, R.J., and Adams, P.D. (2008). Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr D Biol Crystallogr* 64, 61-69.
- Wu, X., Parast, A.B., Richardson, B.A., Nduati, R., John-Stewart, G., Mbori-Ngacha, D., Rainwater, S.M., and Overbaugh, J. (2006). Neutralization escape variants of human immunodeficiency virus type 1 are transmitted from mother to infant. *J Virol* 80, 835-844.
- Wu, X., Yang, Z.Y., Li, Y., Hogenkorp, 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.
- Xiang, Z., and Honig, B. (2001). Extending the accuracy limits of prediction for side-chain conformations. *J Mol Biol* 311, 42 1-430.
- Zhang, Y., and Skolnick, J. (2005). TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res* 33, 2302-2309.

# LIST OF SUPPLEMENTAL FIGURES, TABLES AND MOVIE

- 1 D s , s c u d p - u n HIV- d
- S u i f p e- IV- E v e o Fgu 2
- 3 HIV-1 En type a c o b s nd  
de , d o F u 3.
- C w p gly o , r g h x mpl s d M-g o p  
q , F e 4.
- C w d per ly : ra v o d m e u y m n lys  
l o r
- 6 cov o h HIV-1 E v m urf , a y c mp s o , a  
x p ib n f g yc x of n ay d E v elat d  
6.
- 7 ne fa dy VRC01 y ys a r ure, uc e- i n  
s u , u d i , r to F .
- 1 s y erms, to ma , u d an m a i r
- 2 HIV-1-Env glyc prot n PDB c rd e lys s e d gu
- 3 C y aph a a o f em a , r e gu
- 4 M z l sp c r r d  
Fgu 1 2.
- 5 yc c mpost E v d t b y pr , ea o Fgr S3
- Ta l S6. Av g numbe f yc s v r gw s mp d v h  
j o y 5 0 ns s mu ns, d o Fig 6.
- M v S1 5 s m dynam s imula i fa BG 0 SOSIP- n-9 m d , elat d o  
Fg 6.



## SUPPLEMENTAL FIGURE LEGENDS

**Figure S1. Design, production, and structures of diverse pre-fusion HIV-1 trimers, related to Figure 1.**

(A) Model of BG505 SOSIP (4TVP) with VRC01 superimposed (from 4LSS) showing (B) designed disulfides between VRC01 antibody and gp120 from model. (C) Superdex 200 gel filtration of covalently linked VRC01 -gp140 BG505 IP complex, (D) SDSPAGE analysis of VRC01-gp140 complex, (a) BG505 IP R6.664.G459C/VRC01H(A60CHC)L (b) BG505 IP R6.664.G459C/VRC01H(R61C)L. (E) Characteristics of strains tested and expression yields co-expressed with VRC01 Fab H(A60CHC)L in GnTI<sup>-</sup> cells. (F) Phylogenetic analysis of a group of Envs tested for expression highlighting the strains from which crystal structures are determined in this study. (G) Non-reduced and (H) reduced SDSPAGE analysis of the expressed complexes post Ni<sup>2+</sup> and streptactin purification. (I) Negative stain-electron microscopy of purified HIV-1 gp140s glycoproteins. The proteins were highly homogenous and 2D particle averages are shown. (J) Octet-based antigenic characterization of VRC01 -HIV-1 covalent trimers. (K) Superdex 200 gel filtration chromatograms showing the purification of X1 193.c1-scFv VRC01 complex and the purification of the X1 193.c1-scFv VRC01-PGT122-35O22 ternary complex. (L) Schematic of Man<sub>9</sub>GlcNAc<sub>2</sub> glycan antenna showing the glycan numbering used in the coordinate files. (M) Representative glycan *Fo-Fc* maps at 1 Å after protein-only simulated annealing refinement for gp120 outer domain (left) and V2 apex region (right). (N) Top and (O) side views of X1 193.c1 structure with two protein protomers shown in surface display and one protomer in ribbon display, with PGT122 (blue), VRC01 (black) and 35O22 (gold) and showing gp120 (light gray) and gp41 (dark gray) and glycans (green).

**Figure S2. Structural details of three pre-fusion HIV-1 Envs, related to Figure 2.**

(A) Comparison of pre-fusion HIV-1 gp140 protein structures from clades A, B and G, showing regional differences in protein structure and structural features of HIV-1 gp41 in JR-FL and X11 93 .c 1 structures (shown with *2Fo-Fc* electron density). The X11 93 .c 1 fusion peptide residues 514-521 are observed to form a short alpha-helix of the neighboring protomer via the fusion peptide L519 side chain, G516 main chain and the M515 residue, where the L519 side chain is positioned between the neighboring protomer side chains N644, Q647, E648 and N651. (B) Top view and (C) side view of X11 93 .c1 gp140 trimer with protein shown with B-factor putty and glycans shown in stick, colored according to B-factors. (D) Sequence alignment of the 3 gp140s structurally characterized showing HIV-1 pre-fusion core and variable regions highlighted in black and red respectively. (E) Antibody binding footprints for PGT122 and 35O22 in the context of epitopes from the three HIV-1 Env trimers. VRC01 (black) and 35O22 (gold) and showing gp120 (light gray) and gp41 (dark gray) and glycans (green). (F) Representative X11 93.c1 glycan *Fo-Fc* maps at 1 Å after protein-only simulated annealing refinement showing gp120 V1/V2 apex glycans.

**Figure S3. HIV-1 Env glycans: type, interaction, and correlation between nearest glycan and order, related to Figure 3.**

(A) Hydrophilic Interaction Liquid Chromatography-Ultra Performance Liquid Chromatography (HILIC-UPLC) retention time profile for released glycans from the three HIV-1 Env trimer glycoproteins. Data allows quantitative assessment of the relative proportions of glycans on the protein samples. (B-D) Glycan-glycan interactome displays for the 3 HIV-1 gp140 SOSIP trimer

glycan shields. Connections are indicated by lines where inter-glycan distances are  $< 7 \text{ \AA}$  and the thickness is proportional to the number of interatomic contacts. **(E)** *Fo-Fc* maps at 1  $\text{\AA}$  after protein-only simulated annealing refinement matched to the structures shown in Figure 3C. **(F-H)** Nearest-neighbor analysis reveals that there is no correlation between closest glycan distance and the number of refined saccharide units in the three crystal structures. Red circles correspond to glycans in close contact with antibody.

**Figure S4. Crowded and dispersed glycans: definition, crystallographic examples, and M-group frequencies, related to Figure 4.**

**(A)** Scatter plots showing how the number of ordered saccharide units in x-ray structures from BG505 (clade A), JR-FL (clade B) and x1 193.c1 (clade G) correlates with the number of sequons within some radial shell. A Pearson correlation is used to determine all correlations. The number of active sequons for BG505, JR-FL and x1 193.c1 is 21, 23, and 25 respectively. **(B)** Scatter plot showing optimal partitioning of crowded and dispersed residues at a radius of  $50 \text{ \AA}$  using all 3 HIV-1 gp 140 trimer structures. **(C)** Application of Fisher Exact test to determine the twenty lowest *p*-values for the combined data set at a radius of  $50 \text{ \AA}$  (residues in contact with antibody are not considered here). **(D)** Molecular features, observed crystallographically for crowded glycan N413 and dispersed glycan N61 1 (left side). Observed surrounding inter-glycan interactions for crowded glycan N413 and dispersed glycan N61 1 showing a  $50 \text{ \AA}$  sphere radius and interfaces between contacting glycans highlighted with red dashed lines (right side). **(E)** Top, M-group glycan frequency for most frequent glycans not found in crystallized strains. Bottom, M-group distribution of the number of glycan neighbors within a radial cutoff of  $50 \text{ \AA}$  around each select glycan, obtained by threading each sequence with BG505 (blue), JR-FL

(magenta) and X11 93.C1 (green) structures. Colored circles indicate medians, thick grey lines indicate interquartile range, and thin grey indicate 95% range around the median. Glycans with 15 or more median neighbors were classified as crowded based on Figure 4.

**Figure S5. Crowded and dispersed glycans: natural variation and molecular dynamics analysis, related to Figure 5.**

(A) Glycan frequency difference between subtypes: Glycan frequencies within each subtype are shown for those glycans that had at least 10% glycan frequency in any subtype and were not in hypervariable regions ( $n = 38$ ). The glycans are colored green or orange depending on whether they were classified as crowded and disperse glycans, respectively, for most of the total 2,847 M-group sequences threaded using the X1 193.c1 crystal structure (Methods). The glycan

SciPy), and significance thresholds used were  $q\text{-value} < 0.1$  (using the Python package from <https://github.com/nfusi/qvalue>). Below each glycan for a given subtype, the letters indicate the subtypes that had significantly different glycan frequencies. (B) Schematic illustration depicting torsion angle degrees of freedom for a Man-9 glycan. (C) GlcNAc-GlcNAc-Bma torsional angles belonging to dispersed and crowded glycans show similar distributions. (D) Distribution of GlcNAc-GlcNAc-Bma angle (angle formed from geometric centers of the first three saccharide units of a glycan) for the Man-5, Man-7 and Man-9 simulations. (E) Distributions of inter-glycan contacts (10 or more atoms within  $5.0 \text{ \AA}$  of another glycan) for the Man-5, Man-7 and Man-9 simulations. A Mann-Whitney test was used to determine the medians from both distributions were statistically different. (F) Distribution of GlcNAc-GlcNAc-b-mannose angles defined as the angle formed between the geometric center of the first three saccharide units of a

glycan. The distribution is computed over a 500 ns simulation of a Man-5 model. Distributions for crowded and dispersed glycans appear on the left and right respectively with the angles for the first three saccharide units in the BG505 crystal structures shown in black thatched bars. The red dotted line indicates the average angle (see Figure S5D). **(G)** Inter-glycan interactions from the same 500 ns simulation of the Man-5 model. An inter-glycan interaction is defined as occurring when 10 or more atoms from one glycan are within 5.0 Å of another glycan. Distributions for crowded and dispersed glycans appear on the left and right respectively.

**Figure S6. Glycan coverage of the HIV-1 Env trimer surface, glycan array composition, and examples of antibody recognition of glycan in the context of glycan array and Env, related to Figure 6.**

**(A)** Surface representation of BG505 used in the MD simulation with group M protein sequence conservation colored from white to red according to sequence variability. **(B-D)** Composites of 500 ns for BG505 SOSIP gp140 Man-5, Man-7 or Man-9 MD simulations respectively showing the degree of glycan coverage of the protein surface (protein surface colored as in (A)). Glycans are shown in green with an isosurface contour level at 20% of the maximum density. **(E)** Schematic representation of the 40 N-glycans printed on NHS coated microarray glass slides. **(F, G)** Glycan array binding profiles for antibodies PGT 122 and 35O22. **(H, I)** Electrostatic surface potential representation of antibodies PGT122 and 35O22 bound to glycans.

**Figure S7. Glycan interactions of antibody VRC01 by crystal structure, structure-function studies, and molecular dynamics, related to Figures 6 and 7.**

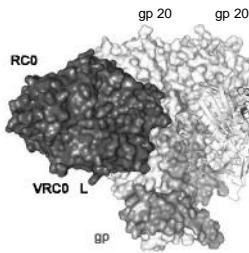
(A) Details of VRC01 paratope and (B) epitope contact surfaces in context of glycosylated HIV-1 X1 193.c1 trimer. VRC01 heavy chain-gp120 protein contacts are shown in yellow and orange (neighboring protomer) and VRC01 heavy chain contacts to glycans are shown in pink. VRC01 light chain-gp 120 protein contacts are shown in blue and contacts to N276 glycan are shown in red. Glycans are depicted in green. (C) 3BNC1 17 light chain mutants were equally potent as mature antibodies in the context of T278A viruses but predicted N276 glycan contact residues were important for neutralization of wild-type viruses. (D) Furthermore VH 1-2\*02 class antibodies VRC01, VRC20, 3BNC60 and 3BNC1 17 with germline-reverted light chains (HgL) were similarly effective as mature antibodies in neutralizing T278A viruses, but were substantially less potent in the context of wild-type viruses with glycan N276 present. (E) VRC01HgL mutants with mature light chain amino acids contacting the N276 glycan added show gain of neutralization function. (F) Relationship between glycan array recognition and crowded and dispersed glycans. (G) Fraction of buried surface area of glycan in antibody epitopes (PDB IDs CH103: 4JAN, PGT135: 4JM2, VRC13: 4YDJ, HJ16: 4YE4, PG9: 3U2S, VRC01: X1193.c1 structure reported here (5FYJ), PGT122: 4TVP, 35O22: 4TVP, 8ANC195: 4P9H, b12: 2NY7, PGT151: 4NUG, CD4: 1GC1). (H) Average HIV-1 Env glycan atoms overlapping broadly neutralizing antibodies and CD4 molecular volumes over the 500 ns trajectory for the three MD simulations.

#### **Movie S1. 500 ns molecular dynamics simulation of a BG505 SOSIP-Man-9 model**

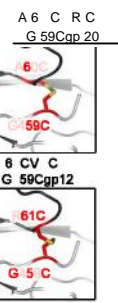
Each glycan sequon on the BG505 SOSIP.664 trimer was modeled with a Man-9 glycan antenna. The simulation was created with ACEMD (see Extended Experimental Methods) and rendered in

PyMOL with 0.1 ns frames. The simulation revealed glycan flexibility occluding the surface of the prefusion HIV-1 Env trimer and was used to compute data found in Figure 6 and Table S4.

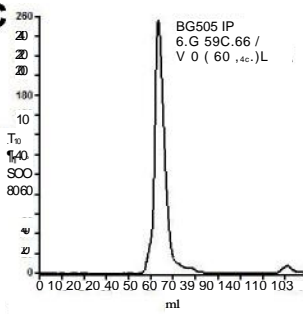
A



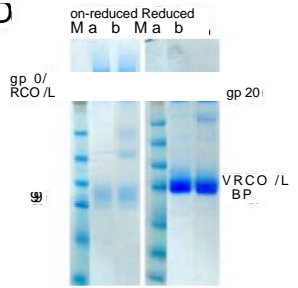
B



C



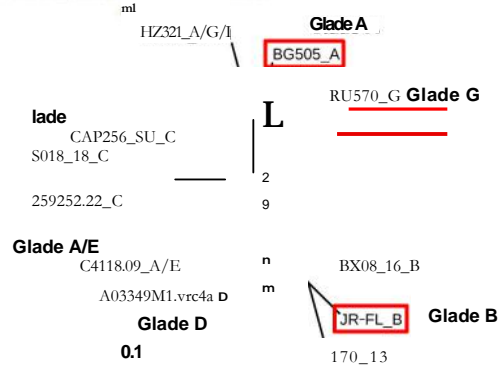
D



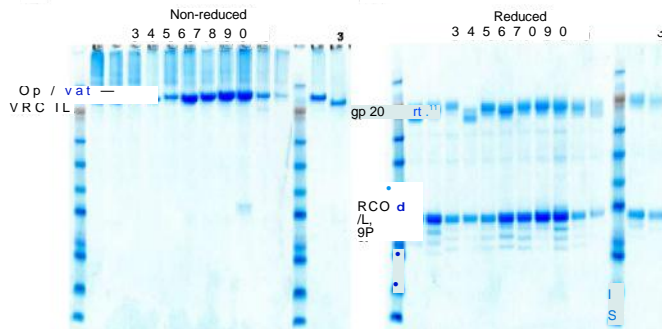
E

D -	P GE	lane	IG.H1 G yo Cl	Mu s	Org /s la a	Prot n Ge b y Id access no	Imo/L number
2	JR- L	B	o IP 59C R 66	US /198	2.8	80560	
2	G 0	A	o IP, 9 R6 68	Kenya/2006	2.9	6 8	
3	WI	B	IP 59C R6 5	U / 0	.6	W6	
	C 8.0	A/	o I C R6	ha land/2		Y78535	
5	C 25 U	C	o I 59C R	5 South f ca/2 5 1.7		GW271 3	
6	o 8_8	C	o IP, 59C R6 66 outh Afr ca/2 8 3.3			N/A	
7	X 93.01	G	o IP 59C R6 5	pa n/2	2.	CJ68896	
8	2592.2	C	o IP C R6 685	Ind ail 9	3.2	ABL67 9	
9	X08_		o IP 5 C R6 6	ance/ 92	3.6	95 53	
	o 9M .vrc a D	o IP 5 C R6	Uganda/ 9			UI2388	
	U7	o I , 9C 668	u a/2 7		.7	BN31b/	
2	3	/G/I 0 IP 5 C 6 6 5	are/ 7			P05	
3	BG	A	I C R	Kenya/ <2.6	0:13	B 6 5	

F

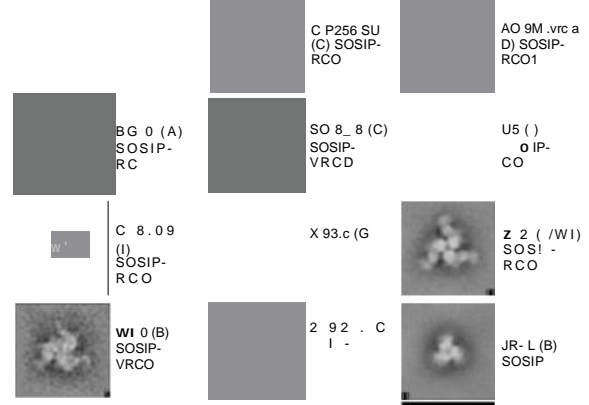


G



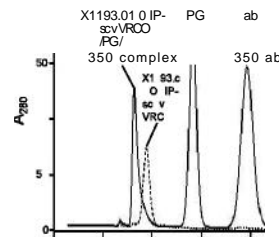
H

I



J

b undt VRC0I Glade PG 4	VRC26.09	PG9	PG	PG 28 3 022
J - L				
BG505	B			
W0	A/			
C 8.9	C			
CAP256 SU	C			
SO 8_8	C			
X 93.c	G			
2 925 .2	C			
BX08_16	B			
033 9M Arm aD G	nd	nd	nd	nd
RU5	nd	nd	nd	nd
Z32	nd	nd	nd	nd



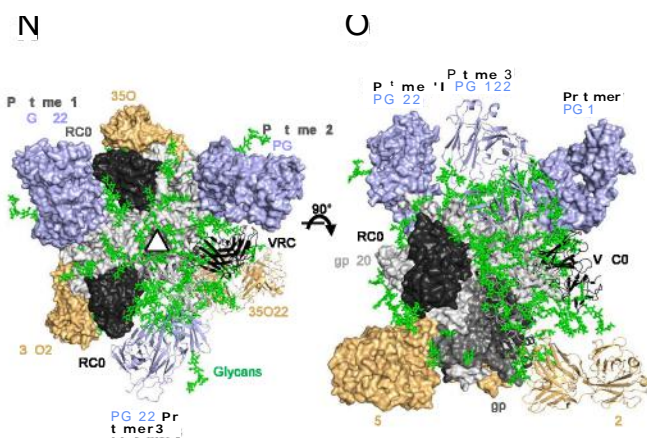
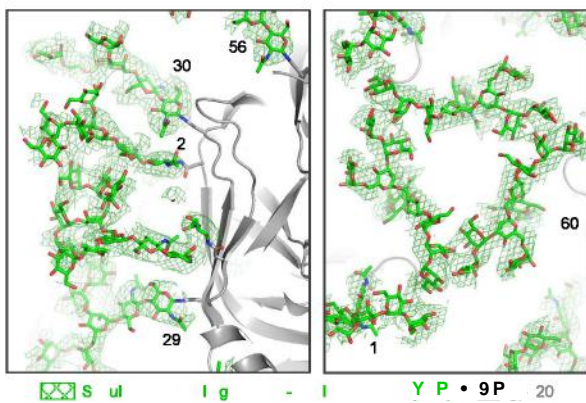


Man ManC ManD1  
1.1 t  
anA anD2

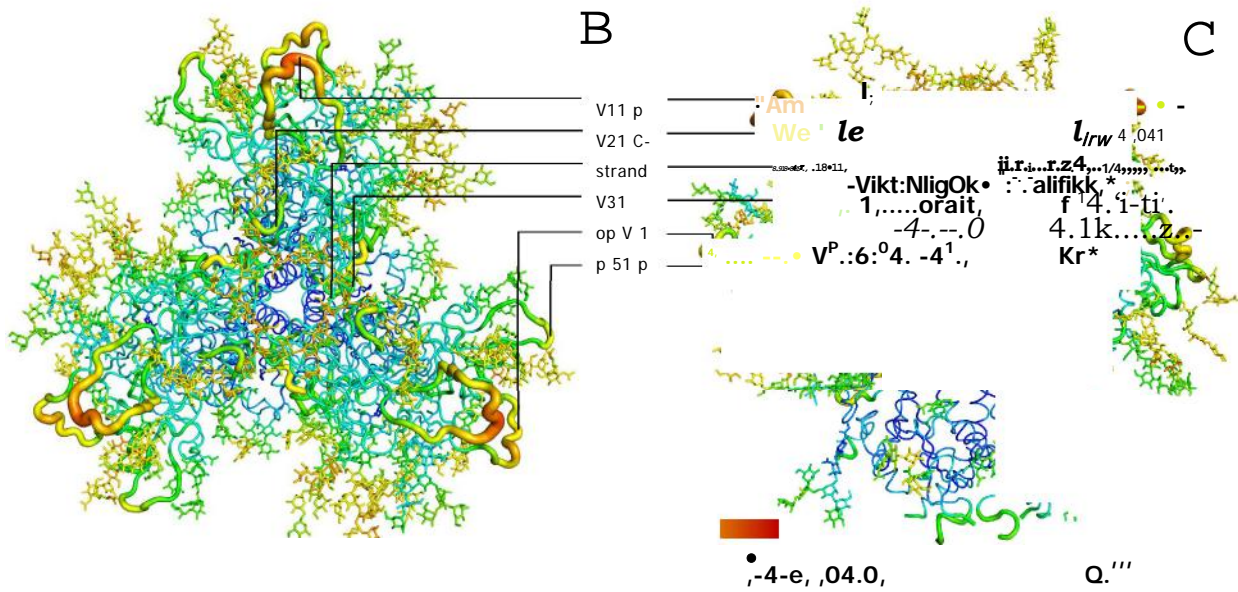
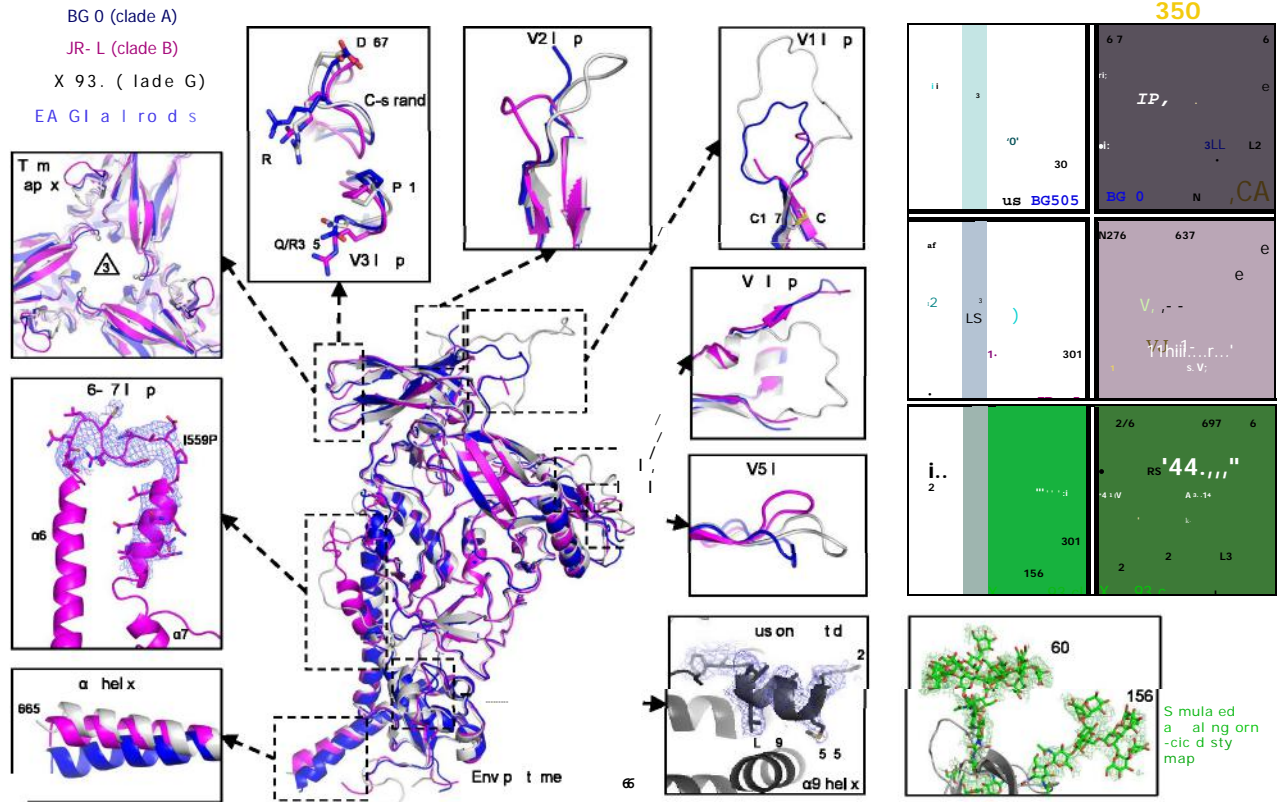
I-1.11 D1

GI NA GI NA an3  
2  
■ 0 M a n \ f i a l - l • / D 3  
Na Nyl Ma ose Glu osamine  
m l

lei l /a\*171—di,t D2  
ManB anD3



g S . D s n , a s u t v s - on IV- m s l a d o .  
 A M d I B 5 5 O S I ( P R C O u p e m p s ( f m L S n B d n d I d  
 R C a b y g 20 m d I . C S u r e 2 g I I a f c v a l t l y I V C -  
 gp BG I c I D) D P A a a l y C -gp m I (a) BG5 I  
 6.6 .G 59 / R C O (A 0 L b B G 50 I P R .66 .G 9 C N C O 6 L . E o  
 s I p I s - x p d V R b 0 L G I - I I . 1  
 I s m d s G) - a S D S P A G I t h  
 x d c m 1 p s t \* d s u a I g - 1 n m p y p u  
 1 - I p s l y 2 D t 1 a v r . (J  
 0 - s d n g c z R C - 1 - v a l m s . K S u 2 0 g 1 I a  
 r o a g r a s w n g u a f X 93 . - v V R C o m p l a d e p u c a o n o e  
 X 9 . c - s v V R C O - P G - 5022 n y c m I x . (L) S e m a f M n , G I g l a  
 o I m b g h c I s . M R p r g l c - c y  
 p t - o 1 m 1 n I m p m I d p g . N  
 a d 0 s d s X 9 . s t u u t p n m s s w s u s l y a  
 m e r d s l a y , P G l u e , R C O I c a 50 I s w n g g 20 I  
 p d 1 c s e e .



gp 1 a6-a7  
loop

allet<sub>0</sub>  
-1V  
et  
.t'44p'

AGH  
90° Twptms  
display d  
cla ty

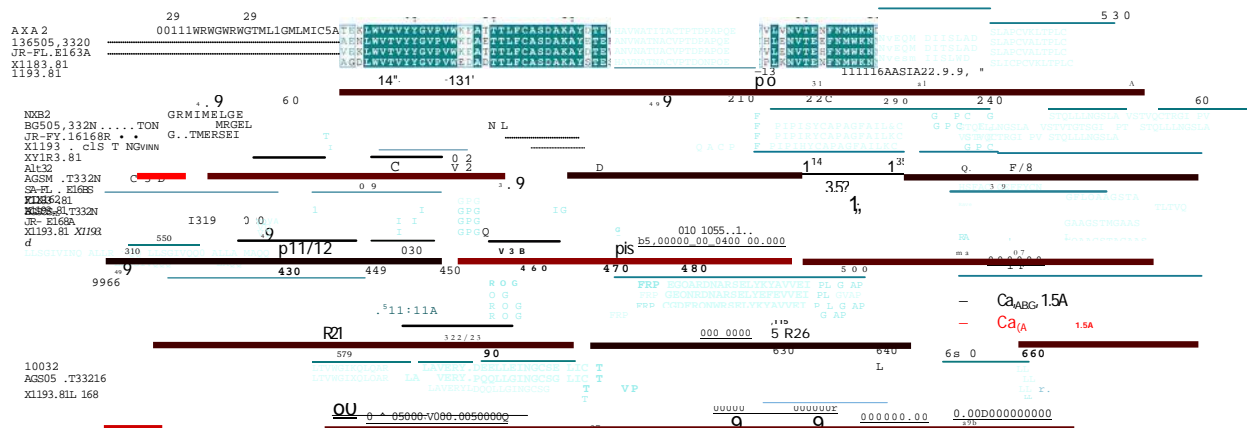
9170.

B-fa for sale ( $A^2$ )

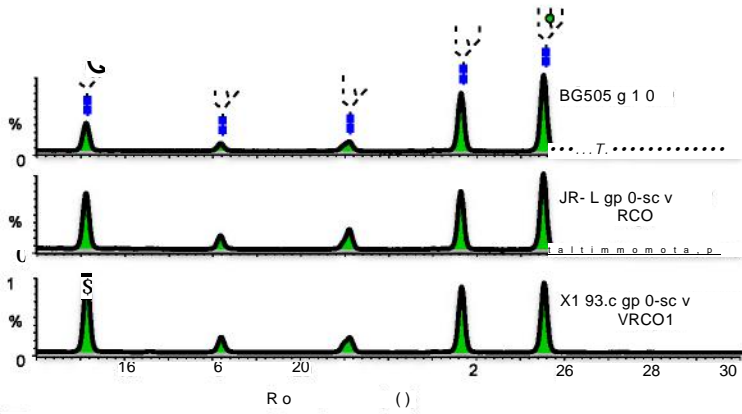
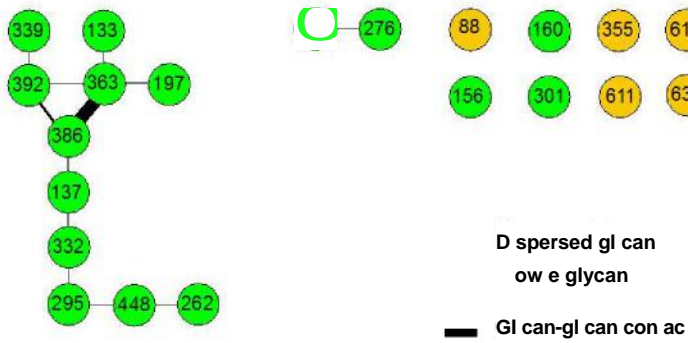
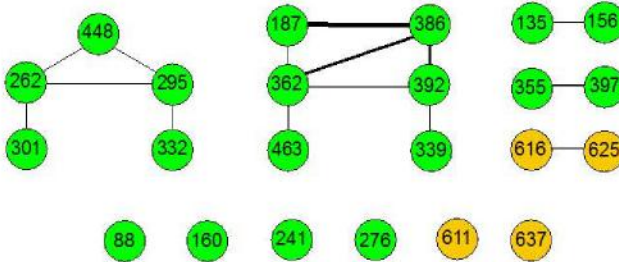
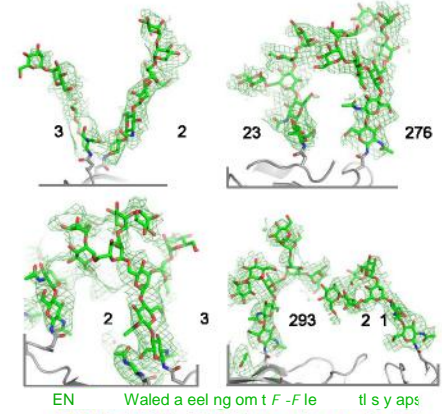
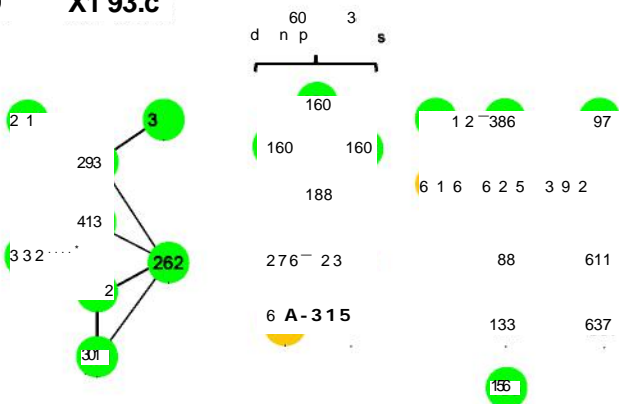
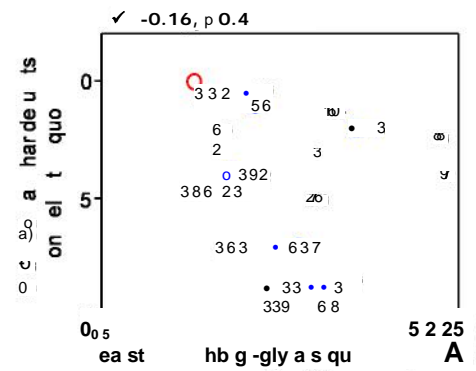
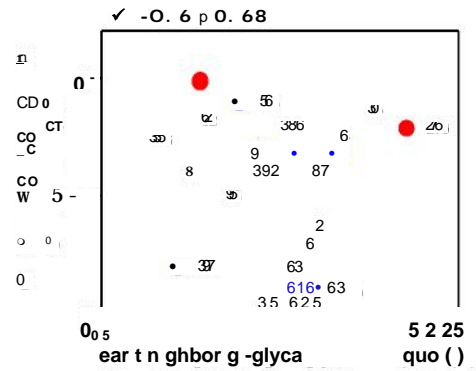
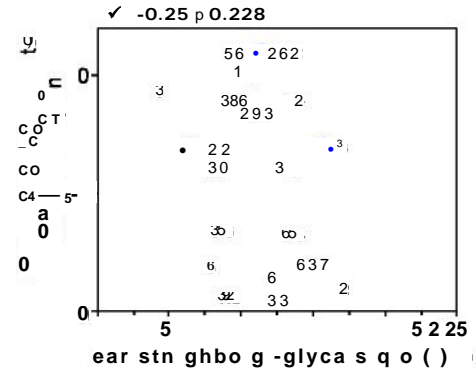
Mir  
X 93.c

20

250



**gur 2. uc l d a l s of hr pr -f s o IV-1 vs, el d o g 2.**  
**(A)** Corn a on o e- u on IV-1 gp1 0 p e n u c u e m Glade A B and G sh w re al  
d p o n s e If t es IV- gp JR- La d X 193. s  
(s ow w t 2 o- cele y). X 93. s - 2 b ma  
s l - l c ng d cula a9 l t gh o ng r m a s  
p L 9s h G m nc a M u h L 9 c s p s d  
a l x g g p o s c s Q 7 8 a . **B) o v C)**  
s d f X 3. p m hB- f o u y d ly sh s  
l d d B-fa .(D)S qu c l g m t f 3g ss lly z h  
l - - us a l o n s g l h d bl sp ct v ly. ( ) A y  
oot s fo PG 22 a d 3 022 n h c n x p o es om ee IV- s. RCO ( lac )  
a d 350 l a s w 20 l ray a d a) a l s g e). ( )  
Rep s a e X 93.c gly a Fo-Fc ps a of r prote - ly s mul d a al g ref sh w g  
g 20 / 2 ex glyc s.

**B BG505****C JR-FL****D X1 93.c****Clade A****G****Clade B****H****Clade G**

gure S3. IV-1  
ela o gu 3.

v gl a s: p

r o,

co l o b we

a est gl c

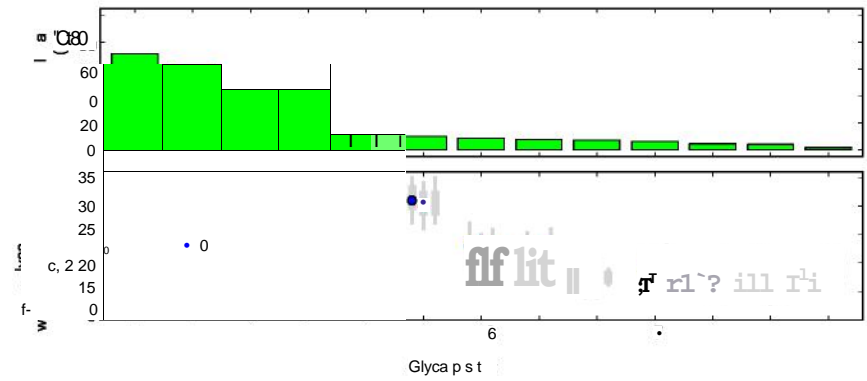
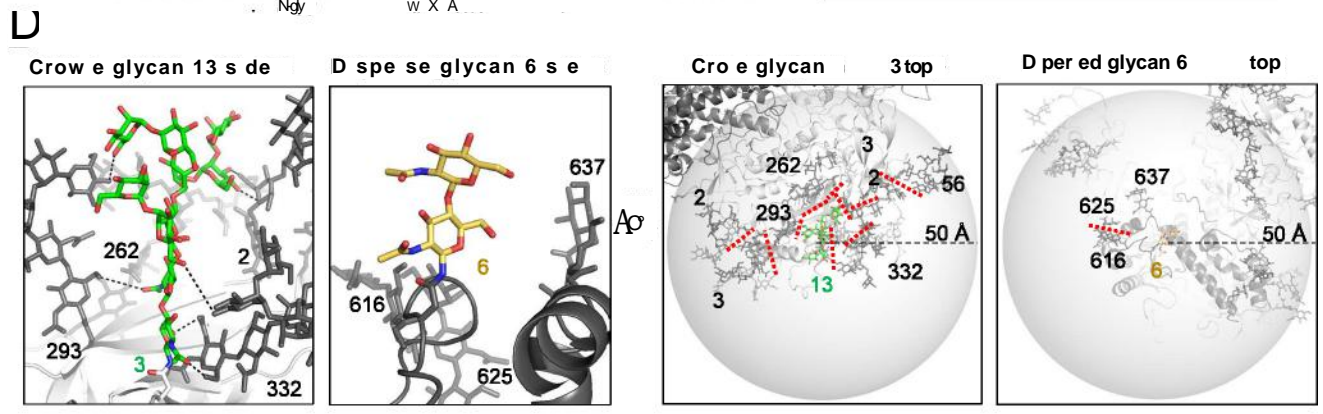
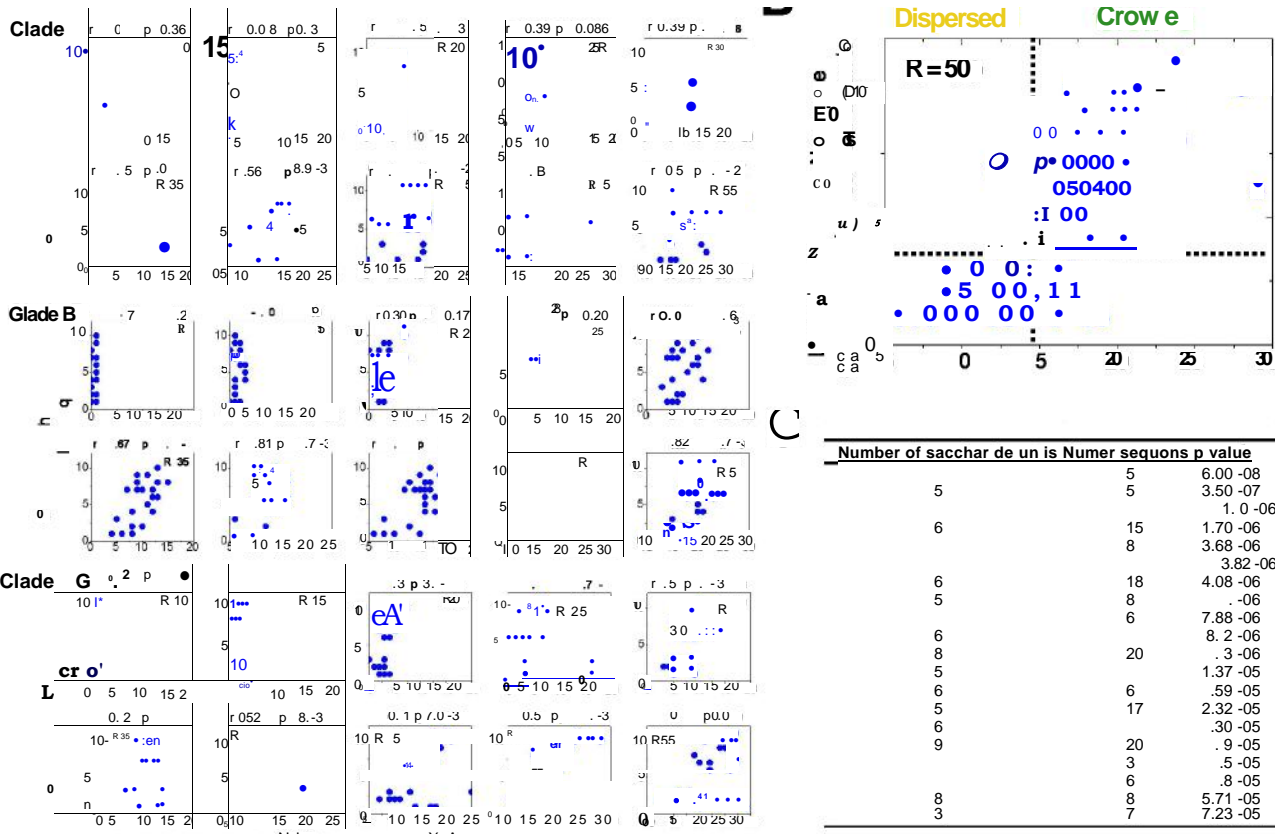
o de

(A) p l l o L u  
r m p l l gl  
q a la  
a t m pl 3 1 - g  
re r-gly d s s ar < 7 Å  
. F-m p t1cra  
u 3. - N  
l s u b  
p d gl a s l a

g y- l P L  
om t l -1  
p l  
0 SOSIP gl s ld .  
k s p  
- 1 l l g f m  
- g b a ly v ls  
a t  
a b dy.

g y( ILI - PLC)  
gl p . D all ws  
a 1 . (B-D) Gl a- l ca  
on a d d b l  
l m o  
m h s  
l n l  
c l . R l



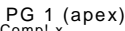




gur S4. Crow a sper glyca d f , ys ll g ph xam l s a M-g u  
 fr qu nc s la o gu 4.  
 ) S <sup>131</sup> d a - u BG5 5  
 (ld A J-L Id Ba 93. l G l u s u m a l  
 s ell. A P a s la d m ll la s. m c qu BG 05  
 JR- L a d 93.c 3 2 e l. (B S l n m l g c  
 an sp u 50A g113 l- p 0 m u e. (C A l o  
 ac m t w n y l s p- al u s o e m e se s 50 A  
 ( s es w a y e n t on d ). (D) M l la fe b v  
 s lb ap cally f w d l 3 sp sed l 6 (lef s ). 0 o n  
 -gl c c o s f c w gl ca 3 an d p s glyc 6 s o g 50 A s p e d u  
 f b o c g gly g l g ed re as l n s r g s . ( ) op -group  
 gl e m s gl an u s ll z a s. B m M-g oup s bu on f  
 mb r gl b rs th a l 50 A a I gl  
 h d h q e h BG505 (bl JR- L ma a a d X 93.0 ( s. C l  
 l m h ck l q a l a a h c 9 % a  
 e . Gly a w m hb w l f w ba g

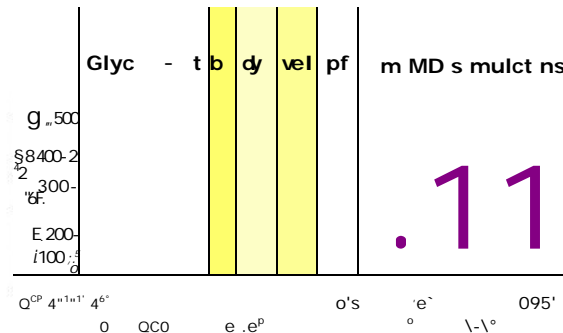
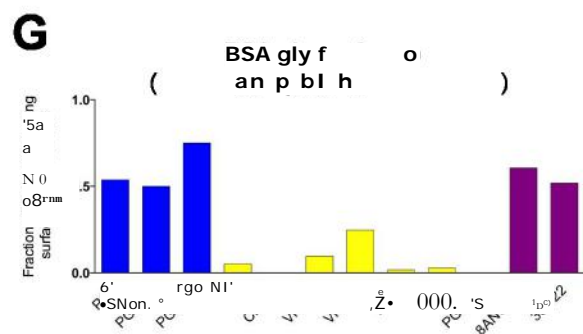
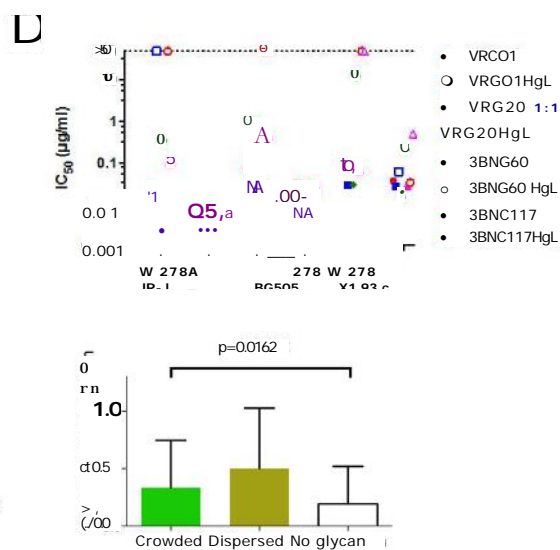
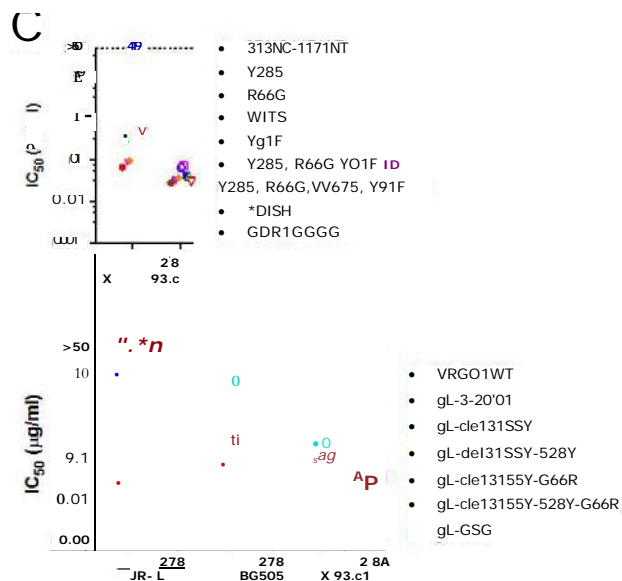
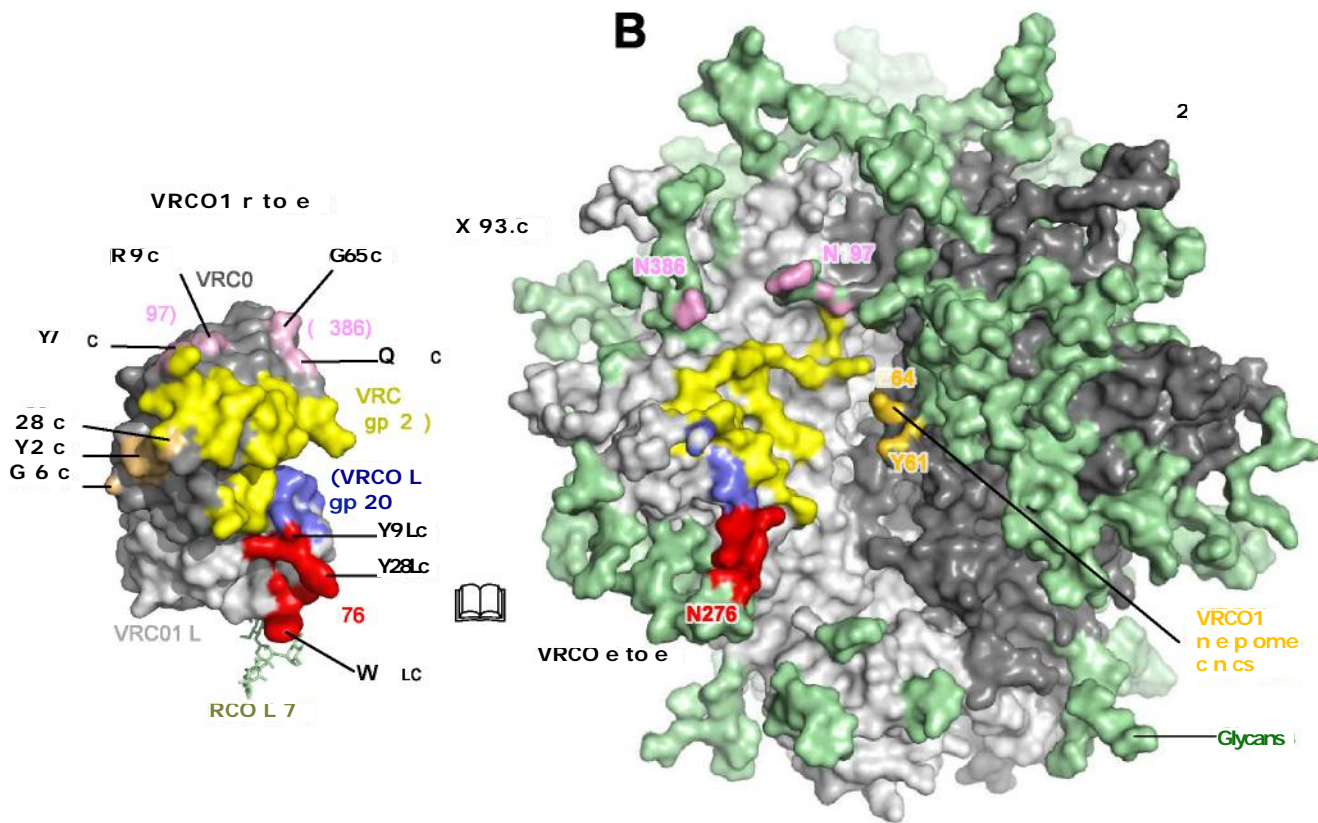


ge S5. Cow e an      spe glyc      a al      d mol ula      m      l      , el  
 to gu e5.  
 ( ) l      q      p      l an      q      a      p a s o      o  
 In      t d      l      0% gl      q      d      t      l      d      d d  
 38).      l a s      l      d      d      t      l      d      d d  
 d      gl a      l      m      al 2 8 7 - u s q      s      a      s      X 93.  
 c      l s c      ( t      ).      l      q      c e s      n      y s w      m p      n      s r' s x a c  
 s      p l      S c )      s g c c      s l s u s d      q - v l      0.      g      s  
 k g f o m p s //      . c m / s / q l u ). B l o      l c      s b      l      s      c a e e  
 s b y p      a      d s      l y d      l y      q u c s. ( B S      m      l l u s      o n      p      l  
 e s f e e      o a      n - 9 l y      . ( C G l      - G l      c - B a o s      l a l      l o g      o s p      a  
 c      d      l y      s s w s l a d s      s. ( D D s      f G l      - G l      - B      l (      l      m  
 g o      c c      s o      f s      e a h      e      o g l y c ) o r h      - 5,      -      d      n - 9 s m l t o s.  
 ) D      b u      s      - g l y      c      (      m o      m      t      A      o f r g l y c ) r e M a - ,  
 d      - 9 s m l o . A      - W      w      o e      e      s f o b o      s b u t o s  
 c l l y      f      t. ) D s b      l      - l      13 -      g l      d      a l  
 b      c c      h      a      e      s      g l y a.      b      o m p d  
 l a      f      -      l. D s b      c      d      l y c      h l a  
 p c l y      g l      u      B G 505 y a l      s o w  
 black th      b s.      d d o      l e      s      a v g      g l ( s      g u S5D ). ( G ) l      - g l y  
 c      5      s      l      h      - 5      e l.      r - g l y c      r c      s  
 o      g w h e ' l      m o a s f      o e g l y a a w h 5.      l      . D b      w d  
 s      l      a      l      g      p      l -



X 93.6  
-2.0 2.0  
lec ros at c pate al

f b e S. GI g f h IV-1 s f c e l p s , x ple  
 I d S B 505 d l w M o s gur .  
 gp 0 -5 -7 a g s u l . (B-D C m s B 50 SOSI  
 P u ac (p o u a of as ( ). Gl h g h a s s a n  
 ll 20% t x u d .())Sche c p a 0 N-gl ca s p S  
 e m ay gl sl d s. ( , G) Gly a y b g p les b PG 2 d 3 . ( , l)  
 le s su a po e l p a on n bo PG 22 3 22 b u gly .



gu S7. Glyca  
m o l a r m l d o g 6 d 7.  
D l s f RCO p pe B ep pe c n su f s n e f glyc s ylated l - X 93.  
m. RCO h avy cha -gp 20 p ac s a h w y l l ad a g n ghb ng pro ome )  
and RC h avy h n a gly a sae sho p k. VRCO1 l g c a n-gp 2 p s  
a s n n bl a d co tac 2 6 glyca a sh d. Glyca sa d pc d ng n. (C  
3B C 7 l g ca m a s ere quail p as ma u a b d s hc x 2 8 v us s  
p d ce 276 gl c n ac s du sw mp a t f al a no w ld-typ s s. (D  
u the m -2\* 2 cla a b d es RCO RC20 3B C60 n 3B C117 w th ge ml e- e e  
l g ha ( gL mlalye c v s ma u e t bo es ne t al ng 2 8A v r e ,b t ere  
s s a ally l po n co e ow ld- ype ses gly a 2 6 p se . ( RCO gL  
m a s mar l g a n am o a s o ac g th 276 glycan a e sho ga of eu ral za on  
f c . Rela o s p b gly a a ay ec g o a c owd d and d sp rs glycans. G  
c f b e su face a ea glyc bo yep pes (PDB IDs C 03 G 5 2  
C YDJ Y 6 G9 3U2S CO X 9. s u e epo ed e e PG 122 VP 35022  
P 8 95 P9 b 2 2 Y7, CD G ). ( ) e e l - l ca ms ye lapp g b o dl  
e al g nt b d sad CD m lc! volumes ve 500 s rajec ory f e D s m la s.



## SUPPLEMENTAL TABLES

Table S1. Glossary of terms.

Term	Description
35O22	A neutralizing monoclonal antibody that specifically binds to an epitope on the gp120-gp41 interface of HIV-1 Env near the viral membrane. The amino acid sequences of the heavy and light variable regions of the 35O22 mAb have been deposited in GenBank with Accession numbers AIO10972.1 (35O22 V <sub>H</sub> ), AIO1 0980.1 (35O22 V <sub>L</sub> ).
BG505	The BG505 strain of HIV is a Clade A HIV-1 virus isolated from a six-week old HIV-1 infected infant. The amino acid sequence of BG505 Env protein can be found under GenBank with Accession number ABA61516.
CD4	Cluster of differentiation factor 4 polypeptide; a T-cell surface protein that mediates interaction with the MHC class II molecule. CD4 also serves as the primary receptor site for HIV on T-cells during HIV infection. CD4 is known to bind to gp120 from HIV-1 Env. The amino acid sequence of human CD4 is deposited in GenBank with Accession code number P01730.1.
Clade A	HIV-1 clade A is a subtype of HIV-1 common in Africa.
Clade B	HIV-1 clade B is a subtype of HIV-1, which is the dominant form in Europe, the Americas, Japan, Thailand, and Australia.
Clade C	HIV-1 clade C is the subtype of HIV-1, which is the most common in the world. It is the dominant clade in Southern Africa, Eastern Africa, India, Nepal, and parts of China.
Clade G	HIV-1 clade G is a subtype of HIV-1, which has been found in West Africa and central Europe.
Complex glycans	Complex N-linked glycans have N-acetyl glucosamine (GlcNAc) residues at -6 mannose sites. Complex glycans do not contain mannose residues apart from the core structure. Additional saccharides may occur in repeating lactosamine (GlcNAc-), tri- and tetraantennary forms making up the majority of cell surface and secreted N-glycans, and complex glycans commonly terminate with sialic acid units. Additional modifications include the addition of a bisecting GlcNAc unit at the mannosyl core or a fucosyl residue on the innermost GlcNAc.
Crowded and dispersed glycan	“Crowded” or “dispersed” glycans were based on the number of neighboring sequons and the number of ordered saccharide units. By using Fisher’s exact test based on a 2 × 2 contingency table, we determined 15 neighboring sequons (within 50 separation between the two classes.
DS-SOSIP	SOSIP with engineered disulfide 201 C-433C. The engineered disulfide stabilizes the Env trimer in its prefusion closed conformation and reduces the antigenicity to non-broadly neutralizing antibodies such as those against the V3 region.



<b>Epitope</b>	An antigenic determinant defined by particular chemical groups or peptide sequences on a molecule that are antigenic, and have elicited a specific immune response. For example, an epitope is the region of a protein or peptide to which B or T cells respond. An antibody can bind to a particular epitope, such as a surface epitope present on the HIV-1 Env structure.
<b>Evolving glycan shield</b>	Evolutionary dynamics of the glycan shield (of the Human Immunodeficiency Virus) over time leading to addition, removal or change in the position of glycan sequons within the viral envelope gene.
<b>Fab</b>	The fragment antigen-binding (Fab) fragment is the region on an antibody that binds to antigens. It is composed of a constant and a variable domain of the antibody heavy and the light chains. The variable domain contains the paratope (region of the antibody that binds antigen) and the complementarity-determining regions (CDRs).
<b>Glycan array</b>	Glycan arrays contain carbohydrate compounds that can be screened with ligands to define carbohydrate specificity and identify binding partners.
<b>Glycoprotein (gp)</b>	A protein containing oligosaccharide chains (glycans) covalently attached to polypeptide side-chains. The carbohydrate chain is attached to the protein during co-translational or post-translational modification.
<b>HEK293S GnTI<sup>-</sup> cells</b>	A human fetal embryonic kidney cell line transformed with adenovirus 5 DNA N-acetylglucosaminyltransferase I (GnTI) activity and therefore lack complex N-glycans. line was established by methanesulfonate mutagenesis followed by Ricin selection and has been used to overexpress a wide variety of mammalian proteins.
<b>High-mannose Glycan</b>	High-mannose glycans contain numerous unmodified terminal mannose sugars. These glycans typically contain between five and nine mannose residues (Man-5, Man-6, Man-7, Man-8 and Man-9) attached to a GlcNAc <sub>2</sub> core. The name abbreviations are indicative of the total number of mannose residues in the structure.
<b>HIV envelope glycoprotein (Env)</b>	The HIV envelope glycoprotein is synthesized as a gp160 precursor of 845-870 amino acids. Individual gp160 polypeptides form a homotrimer and are glycosylated within the endoplasmic reticulum, processed in the Golgi apparatus and cleaved by a cellular protease near residue 512 to produce gp120 and gp41 subunits.
<b>HIV-1 gp120</b>	A polypeptide spanning HIV-1 Env residues 31-511 and containing most of the external, solvent-exposed domains of the HIV-1 Env trimer. gp120 is heavily N-glycosylated and has a molecular weight of approximately 120 kDa. gp120 binds both to the cellular CD4 receptor and co-receptor (e.g. CCR5 or CXCR4). gp120 also interacts with gp41, assembling into a functional HIV-1 Env trimer.
<b>HIV-1 gp140</b>	A recombinant form of HIV-1 Env that includes gp120 and the gp41 regions, but lacks the gp41 transmembrane and cytosolic domains. HIV-1 gp140 polypeptides trimerize to form a soluble HIV-1 Env trimer.
<b>HIV-1 gp41</b>	gp41 spans approximately HIV-1 Env residues 512-860, including cytosolic, transmembrane, and ectodomains. The gp41 ectodomain (comprising

	approximately HIV-1 Env residues 512-644) can associate with gp120 to form an HIV-1 Env protomer that forms the HIV-1 Env trimer.
Human immunodeficiency virus type 1 (HIV-1)	A retrovirus that causes immunosuppression in humans (HIV disease), and leads to a disease complex known as the acquired immunodeficiency syndrome (AIDS) leading to the development of opportunistic infections. Treatment of HIV-1 with HAART has effectively reduced viral burden and alleviating the effects of HIV-1 infection in infected individuals.
Hybrid glycans	Hybrid glycans are characterized as containing both unsubstituted terminal mannose residues (as are present in high-mannose glycans) and substituted mannose residues with an N-acetylglucosamine linkage (as are present in complex glycans). These GlcNAc sequences added to the N-linked glycan core in hybrid and complex N-glycans are called "antennae".
IgG	Immunoglobulin G (IgG) is a type of antibody composed of four peptide chains: two identical heavy chains and two identical light chains and each IgG contains two individual antigen binding sites. IgG molecules are produced from and released by plasma B cells.
JR-FL	The JR-FL strain of HIV, which is a Clade B HIV-1 virus isolated from an HIV-1 infected individual in the USA in 1986. The amino acid sequence of JR-FL Env protein can be found under GenBank accession number AAB05604.
Kink angle	The angle defined by the geometric centers of the pyranose rings from the GlcNAc-GlcNAc- -mannose trisaccharide of a glycan.
M-group	With 'M' for "major", this is by far the most common type of HIV, with more than 90% of HIV/AIDS cases deriving from infection with HIV-1 group M. The M group is subdivided further into clades. A manually curated alignment of all M-group HIV-1 Env sequences can be downloaded from the Los Alamos HIV-1 databases ( <a href="http://www.hiv.lanl.gov/content/index">http://www.hiv.lanl.gov/content/index</a> ).
Molecular dynamics	Molecular dynamics (MD) is a computer simulation method based on Newton's classical equations of motion for studying the dynamics of atoms. The equations of motion are solved numerically allowing the dynamic evolution of the system to be followed.
N-linked glycan sequon	An N-linked glycan sequon is an Asn-X-Ser or Asn-X-Thr sequence, where X is any amino acid except proline and the glycan may be composed of N-acetyl galactosamine, galactose, neuraminic acid, N-acetylglucosamine, fucose, mannose, and other monosaccharides.
Neutralizing antibody	An antibody which reduces the infectious titer of an infectious agent by binding to a specific antigen on the infectious agent. In some examples the infectious agent is a virus. In some examples, an antibody that is specific for HIV-1 Env neutralizes the infectious titer of HIV. A "broadly neutralizing antibody" is an antibody that binds to and inhibits the function of related antigens, such as HIV-1 Env from more than one clade or a high percentage of the many types of HIV in circulation.
PGT122	A neutralizing monoclonal antibody that specifically bind to the glycan-V3 region of HIV-1 Env. The amino acid sequences of the heavy and light variable regions of the PGT122 mAb have been deposited in GenBank with Accession numbers JN201895.1 (PGT122 V <sub>H</sub> ), JN201912.1 (PGT122 V <sub>L</sub> ).

Protein core	A protein core is the interior of a folded protein, which is substantially free of solvent exposure and in general, a protein core is predominately composed of hydrophobic or apolar amino acids.
Root-mean-square deviation (RMSD)	The square root of the arithmetic mean of the squares of the deviations from the mean. RMSD is used as a way of expressing deviation from the coordinates of a reference three dimensional structure. This number is typically calculated after optimal superposition of two structures.
SOSIP	Mutations incorporated into HIV-1 Env ectodomain that facilitates the production of cleaved, soluble Env trimers stabilized with an engineered disulfide bond (501 C-605C) and a helix-disrupting mutation I559P.
STD NMR	The Saturation Transfer Difference Spectroscopy Nuclear magnetic resonance (STD NMR) experiment is a spectroscopic technique to study the interactions, in solution, between a large molecule (receptor) and a medium-small sized molecule (ligand), and is based on the Nuclear Overhauser effect (NOE) and the observation and analysis of the resonances of the ligand protons.
Vaccine	A pharmaceutical composition that can elicit a prophylactic or therapeutic immune response in a subject. It is administered with the aim of inducing a useful or protective immune response. Typically, a vaccine triggers an antigen-specific immune response to a specific antigen or pathogen, for example a viral or bacterial glycoprotein.
VRC01	A neutralizing monoclonal antibody that specifically binds to the CD4 binding site on HIV-1 Env and can inhibit HIV-1 infection of target cells. The amino acid sequences of the heavy and light variable regions of the VRC01 mAb are known and have been deposited in GenBank with Accession numbers ADF47181.1 (VRC01 V <sub>H</sub> ) and ADF47184.1 (VRC01 V <sub>L</sub> ).
X1193.c1	The X1193.c1 strain of HIV, which is a Clade G HIV-1 virus isolated from an HIV-1 infected individual in Spain in 2000. The amino acid sequence of X1 193.c1 Env protein can be found under GenBank with Accession number AAB05604.

---

**Table S2. HIV-1-Env glycoprotein PDB coordinate analysis, related to Figures 1 and 2.**

Env	Env construct	PDB ID	Glycan mass (Da)	Protein mass (Da)	Ratio glycan/protein	Reference
Domain	pfam00516 eODmV3	3TYG	2737.2	16461.9	0.17	Pejchal 2011 <i>Science</i>
	ZM109 V1V2	3U2S	1330.6	4306.8	0.31	McLellan 2011 <i>Nature</i>
gp120	HXBc2 gp120	1GC1	2445.1	30763.2	0.08	Kwong 1998 <i>Nature</i>
	93TH057 gp120	4JKP	3117.4	35788.3	0.09	Diskin 2013 <i>J Exp Med</i>
	YU2 gp120	4RQS	4752.1	30455.1	0.16	Kong 2015 <i>Proteins</i>
Trimer	KNH1144 SOSIP	4CC8	0.0	112377.8	0.00	Bartesaghi 2013 <i>NSMB</i>
	BG505 SOSIP	4NCO	23722.3	138631.8	0.17	Julien 2013 <i>Science</i>
	BG505 SOSIP	3J5M	0.0	137683.6	0.00	Lyumkis 2013 <i>Science</i>
	BG505 SOSIP	4TVP	10150.5	60557.8	0.17	Pancera 2014 <i>Nature</i>
	BG505 SOSIP	4ZMJ	4600.1	59931.4	0.08	Kwon 2015 <i>NSMB</i>
	BG505 SOSIP	5A8H	7983.8	95069.4	0.08	Scharf 2015 <i>Cell</i>
	BG505 SOSIP	5ACO	33416.9	179152	0.19	Lee 2015 <i>Structure</i>
	BG505 SOSIP	5C7K	12013.2	60067.6	0.20	Kong 2015 <i>Acta Cryst. D</i>
	BG505 SOSIP	5CEZ	7071.2	62490.9	0.11	Garces 2015 <i>Immunity</i>
	BG505 SOSIP	5CJX	13343.8	154545.9	0.09	Scharf 2015 <i>Cell</i>
	JR-FL deltaCT	5FUU	35960.2	167932.9	0.21	Lee 2016 <i>Science</i>
	X1193.c1 SOSIP	5FYJ	24596.5	66102.9	0.37	This study
	JR-FL SOSIP	5FYK	20300.7	63095.4	0.32	This study
	BG505 SOSIP	5FYL	18171.8	60557.7	0.30	This study

**Table S3. Crystallographic data and refinement statistics, related to Figures 1 and 2.**

HIV-1 gp140 trimer	Glycosylated X1 193.c1 SOSIP.665 scFv VRC01, Fabs PGT122 & 35O22	Glycosylated JR-FL SOSIP.664 scFv VRC01, Fabs PGT122 & 35O22	Glycosylated BG505 SOSIP.664 Fabs PGT122 & 35O22
PDB ID	5FYJ	5FYK	5FYL
Data collection			
Space group	P6 <sub>3</sub>	P6 <sub>3</sub>	P6 <sub>3</sub>
Cell dimensions			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	127.16, 127.16, 313.65	130.78, 130.78, 314.62	129.78, 129.78, 313.06
(°)	90.0, 90.0, 120.0	90.0, 90.0, 120.0	90.0, 90.0, 120.0
Resolution (Å)	50.00–3.10 (3.55–3.41) (3.41–3.29) (3.29–3.19) (3.19– 3.10)	50.00–3.10 (3.91–3.68) (3.68–3.49) (3.49–3.34) (3.34– 3.21) (3.21–3.10)	50.00–3.10 (3.91–3.68) (3.68–3.49) (3.49–3.34) (3.34– 3.21) (3.21–3.10)
Unique reflections	51,684	55,001	53,617
<i>R</i> <sub>sym</sub> or <i>R</i> <sub>merge</sub>	0.118 (0.579) (0.797) (0.969) (1.00)	0.109 (0.380) (0.403) (0.496) (0.558) (0.784)	0.099 (0.475) (0.490) (0.635) (0.698) (0.748)
<i>R</i> <sub>pim</sub>	0.075 (0.376) (0.533) (0.699) (0.842)	0.066 (0.237) (0.267) (0.355) (0.436) (0.714)	0.060 (0.231) (0.246) (0.329) (0.369) (0.434)
<i>I</i> / <i>I</i>	9.05 (2.00) (1.44) (0.95) (0.64)	7.76 (2.19) (1.51) (1.09) (0.80) (0.55)	4.82(2.12) (1.59) (1.16) (0.71) (0.67)
Completeness (%)	99.7 (100) (100) (99.7) (97.1)	86.2 (53.2) (42.1) (32.0) (22.8) (13.7)	85.1 (56.2) (46.3) (38.0) (31.1) (21.6)
Redundancy	6.0 (6.1) (5.9) (5.4) (4.2)	3.7 (2.9) (2.5) (2.1) (1.8) (1.3)	5.0 (4.1) (3.8) (3.4) (3.1) (2.9)
Refinement			
Resolution (Å)	45.0–3.10	45.0–3.10	45.0–3.10
Reflections	42,755	31,882	32,566
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub>	21.4/27.2	25.3/30.9	26.0/30.7
No. atoms			
Protein	13481	13285	11345
Carbohydrate	1935	1565	1388
Ligand/ion	46	-	-
Water	12	-	-
<i>B</i> -factors			
Protein	120	168	122
Ligand/ion	123	-	-
Carbohydrate	159	195	161
Water	34	-	-
R.m.s deviations			
Bond lengths (Å)	0.005	0.006	0.005
Bond angles (°)	1.060	1.081	1.001
Ramachandran statistics			
Allowed (%)	99.5	99.6	99.7
Favored (%)	90.6	89.9	89.9
C <sub>1</sub> carbohydrate geometry (%)	100	100	100
MolProbity all-atoms clashscore	4.38	5.76	5.21
MolProbity score	2.48	2.65	2.60

Table S4. Modelled glycan size and real space correlation coefficients (RSCC) for glycans, related to Figures 1 and 2.

X1193.c1			JR-FL			BG505		
Sequon	Glycan residues	RSCC	Sequon	Glycan residues	RSCC	Sequon	Glycan residues	RSCC
N88	8	0.81	N88	7	0.81	N88	7	0.73
N133	1	0.39	N135	1	0.50	N133	2	0.61
N142	1	0.29	N156	9	0.65	N137	7	0.52
N156	11	0.44	N160	7	0.34	N156	9	0.70
N160	9	0.44	N187	7	0.52	N160	8	0.55
N188	10	0.36	N241	4	0.70	N197	6	0.47
N197	3	0.68	N262	9	0.77	N234	5	0.67
N234	6	0.63	N276	8	0.72	N262	8	0.76
N241	7	0.60	N295	5	0.61	N276	5	0.43
N262	11	0.81	N301	8	0.60	N295	7	0.62
N276	9	0.62	N332	10	0.78	N301	8	0.59
N293	9	0.59	N339	8	0.56	N332	10	0.82
N301	6	0.74	N355	7	0.49	N339	1	0.60
N332	10	0.76	N362	4	0.69	N355	1	0.64
N344	7	0.48	N386	8	0.51	N363	3	0.71
N355	3	0.36	N392	6	0.44	N386	6	0.62
N386	8	0.51	N397	2	0.71	N392	6	0.34
N392	1	0.30	N448	6	0.71	N448	6	0.58
N413	10	0.53	N463	2	0.37	N611	1	0.30
N442	6	0.56	N611	3	0.45	N618	1	0.43
N464A	2	0.04	N616	1	0.54	N637	3	0.29
N611	2	0.71	N625	1	0.59			
N616	2	0.59	N637	1	0.08			
N625	3	0.59						
N637	2	0.71						

Table S5. Glycan composition of Env and antibody constructs, related to Figure S3.

Glycoprotein	Abundance of Man <sub>5-9</sub> GlcNAc <sub>2</sub> (labelled M <sub>5-9</sub> N <sub>2</sub> ) (%)				
	M <sub>5</sub> N <sub>2</sub>	M <sub>6</sub> N <sub>2</sub>	M <sub>7</sub> N <sub>2</sub>	M <sub>8</sub> N <sub>2</sub>	M <sub>9</sub> N <sub>2</sub>
BG505 gp140	15.3	4.1	7.2	31.7	41.7
JR-FL gp140	24.6	5.8	10.3	25.6	33.7
X1193.c1 gp140	30.4	6.2	8.4	26.8	28.3
PGT122 Fab	>99	-	-	-	-
VRC01 scFv	>99	-	-	-	-

[illegible]



[Click here to access/download](#)

[ppleal oiea pa e](#)

[VRC\\_ D\\_HIV- \\_BG505\\_ an-9\\_ 0 6-0 -12.mp4](#)