HHS Public Access

Author manuscript

Nat Struct Mol Biol. Author manuscript; available in PMC 2016 June 21.

Published in final edited form as:

Nat Struct Mol Biol. 2016 January ; 23(1): 81–90. doi:10.1038/nsmb.3144.

Structures of HIV-1-Env V1V2 with broadly neutralizing antibodies reveal commonalities that enable vaccine design

Jason Gorman[1], **Cinque Soto**[1], **Max M. Yang**[1], **Thaddeus M. Davenport**[2], **Miklos Guttman**[2], **Robert T. Bailer**[1], **Michael Chambers**[1], **Gwo-Yu Chuang**[1], **Brandon J. DeKosky**[1,3], **Nicole A. Doria-Rose**[1], **Aliaksandr Druz**[1], **Michael J. Ernandes**[1], **Ivelin S. Georgiev**[1], **Marissa C. Jarosinski**[1], **M. Gordon Joyce**[1], **Thomas M. Lemmin**[4], **Sherman Leung**[1], **Mark K. Louder**[1], **Jonathan R. McDaniel**[3], **Sandeep Narpala**[1], **Marie Pancera**[1], **Jonathan Stuckey**[1], **Xueling Wu**[1], **Yongping Yang**[1], **Baoshan Zhang**[1], **Tongqing Zhou**[1], **NISC Comparative Sequencing Program**[5], **James C. Mullikin**[5], **Ulrich Baxa**[6], **George Georgiou**[3], **Adrian B. McDermott**[1], **Mattia Bonsignori**[7], **Barton F. Haynes**[7], **Penny L. Moore**[8,9,10], **Lynn Morris**[8,9,10], **Kelly K. Lee**[2], **Lawrence Shapiro**[1,11], **John R. Mascola**[1], and **Peter D. Kwong**[1,#]

[1]Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland [2]Department of Medicinal Chemistry, University of Washington, Seattle, Washington [3]Department of Chemical Engineering, University of Texas at Austin, Austin, Texas [4]Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, California [5]NIH Intramural Sequenicng Center (NISC), National Human Geonome Research Institute, National Institutes of Health, Bethesda, Maryland [6]Electron Microscopy Laboratory, Cancer Research Technology Program, Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, Frederick, Maryland [7]Department of Medicine, Duke University Medical Center, Durham, North Carolina [8]Center for HIV and STIs, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa [9]Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa [10]Centre for the AIDS Programme of Research in South Africa

[#]To whom correspondence should be addressed: (PDK), pdkwong@nih.gov.

Accession Codes. Coordinates and structure factors for Fab CH03 with 1VH8-scaffolded V1V2 (strain CAP256-SU) and Fab CH04 with 1VH8-scaffolded V1V2 (strain A244) have been deposited in the Protein Data Bank with PDB IDs 5ESV and 5ESZ. The new next generation sequencing data for donor CH0219 (454 pyrosequencing and Illumina) used in this study has been deposited in the National Center for Biotechnology Information Short Reads Archives (SRA) with study accession number: SRP065493.

Author contributions J.G. headed the determination of the V1V2 bound CH03 and CH04 crystal structures, revertant neutralization strategy, chimeric SOSIP design and assessment and assisted with VRC26 model development; CS headed the next generation sequencing lineage analysis and VRC26 modeling; M.Y assisted with crystallization. T.M.D., M.G., K.L. performed HDX experiments and analyzed the data; R.T.B. and M.K.L. assessed neutralization breadth; S.N., M.C. and A.B.M. performed antigenic analyses; G.Y.C. performed frequentist probability analysis; B.J.D., J.R.M., G.G; X.U. J.M., NISC contributed NGS data; J.G., C.S., M.P., N.D.R., M.J.E., M.C.J. and B.Z. contributed to paratope mapping; A.D. expressed V1V2 scaffold and antibodies; J.G., I.S.G., Y.Y. and S.L. contributed to V1V2 scaffold design and assessment; C.S. T.L. and JG contributed to MDFF analysis; M.P. assisted with chimeric SOSIP design; J.S. assisted with figure conception and design; U.B. performed EM; T.Z., M.G.J. contributed to reverted VRC01 experiments; M.B., B.F.H. contributed CH0219 materials; P. L., L.M. contributed CAP256 materials; J.G., C.S., L.S. and P.D.K. assembled and wrote the paper, on which all principal investigators commented.

Competing financial interests The authors declare no competing financial interests. Reprints permissions information are available at www.nature.com/reprints/index.html.

(CAPRISA), University of KwaZulu-Natal, Congella, South Africa [11]Department of Biochemistry & Molecular Biophysics and Department of Systems Biology, Columbia University, New York

Abstract

Broadly neutralizing antibodies (bNAbs) against HIV-1-Env V1V2 arise in multiple donors. However, atomic-level interactions had only been determined with antibodies from a single donor, making commonalities in recognition uncertain. Here we report the co-crystal structure of V1V2 with antibody CH03 from a second donor and model Env interactions of antibody CAP256-VRC26 from a third. These V1V2-directed bNAbs utilized strand-strand interactions between a protruding antibody loop and a V1V2 strand, but differed in their *N*-glycan recognition. Ontogeny analysis indicated protruding loops to develop early, with glycan interactions maturing over time. Altogether, the multidonor information suggested V1V2-directed bNAbs to form an 'extended class', for which we engineered ontogeny-specific antigens: Env trimers with chimeric V1V2s that interacted with inferred ancestor and intermediate antibodies. The ontogeny-based design of vaccine antigens described here may provide a general means for eliciting antibodies of a desired class.

Antibodies capable of neutralizing a majority of circulating HIV-1 isolates develop in approximately half of those infected with HIV-1 for over five years[1]. Intense interest has focused on these antibodies, as they provide clues to how an effective vaccine might be developed[2,3]. In specific, broadly neutralizing antibodies (bNAbs) –that arise in multiple donors and share common features of Env recognition and B-cell ontogeny– may have utility as vaccine templates, due to the potential for similar antibodies to be elicited by a common immunogen (or common set of immunogens) in the general population[4,5].

An increasing number of such "multidonor" bNAbs have been identified, such as those of the VRC01 class[6,7] (named for the first antibody of the class), which share 'class' features of molecular recognition and B-cell ontogeny[8–11]. This commonality has motivated the development of immunogens, designed to target class-specific features of recognition and to overcome class-specific roadblocks in developmental ontogeny, and success with this strategy has been achieved with immunogens capable of priming the initial stage of VRC01-class development in mouse models[12,13]. This success has heightened the search for other classes of bNAbs with commonalities in recognition and ontogeny. While antibodies against the same supersite of HIV-1 vulnerability often show diverse modes of recognition[4,11,14], bNAbs against the membrane-distal V1V2 apex of pre-fusion closed conformation of HIV-1 Env appear to share a number of characteristics.

Thus far, V1V2-directed bNAbs have been identified in four donors: the CHAVI donor 0219 (CH0219), with bNAbs CH01–CH04[15]; the CAPRISA 256 donor, with bNAbs CAP256-VRC26.01-12[16]; and the IAVI protocol G donors 24 and 84, with the IAVI 24 bNAbs PG9 and PG16[17]; and IAVI 84 bNAbs PGT141–145[18] and PGDM1400–1412[19]. Structures of the ligand-free forms of these antibodies reveal a protruding third heavy chain complementarity determining region (CDR H3), which is anionic, often tyrosine sulfated, and critical for Env interaction[15,16,19–22]. The epitope appears to be quaternary in nature and to include an *N*-

linked glycan at residue 160 along with strand C of V1V2[15,16,18,22–25]. In terms of B-cell ontogeny, approximations of the unmutated common ancestor (UCA) have been inferred for V1V2-directed bNAb lineages from donors CH0219 and CAP256[15,16], which indicate the long anionic CDR H3 to be a product of recombination[15,16]. Initial recognition of UCA (or of V-gene reverted approximations) appears to be restricted to select strains of HIV-1 (e.g. CAP256-SU or ZM233)[15,16,20], to use similar D genes and in some cases related V genes, and to contain similar motifs (e.g. YYD) in the CDR H3[15–18]. Despite this extensive commonality, it has been unclear whether these bNAbs recognize HIV-1 Env through a common molecular mechanism.

The only atomic-level information on the interactions of these V1V2-directed bNAbs with HIV-1 Env derives from scaffolded V1V2-structures of antibodies PG9 and PG16 from donor IAVI 24, where the CDR H3 penetrates to interact in a parallel intermolecular strand association with strand C of V1V2, and *N*-glycans from residues 160 and 156 (or 173) are recognized by the heavy and light chains of the antibody[22,26]. The combination of this atomic-level information and negative-stain electron microscopy (EM) of PG9 complexed to a soluble trimeric Env mimic, BG505 SOSIP.664[27], suggest the quaternary dependency of PG9 to arise from its recognition of glycan N160 from a neighboring protomer[24].

The absence of atomic-level information on the recognition of V1V2-directed bNAbs from other donors raises a number of questions. Do V1V2-directed bNAbs from the other donors use a similar strand-strand association in their recognition? What is the basis of their quaternary specificity: does it arise like PG9 from glycan recognition? And if their molecular features were indeed similar, how would one devise a common set of immunogens to spur their development? Here we set out not only to provide an atomic-level understanding of V1V2-directed bNAb recognition in another donor, but to use this information in the design of class-specific antigens. To facilitate crystallization, we designed trimeric V1V2-scaffolds capable of interacting with quaternary-specific V1V2-directed bNAbs, and determined co-crystal structures with the antigen-binding fragments (Fabs) of bNAbs CH03 and CH04, both from donor CH0219. We used hydrogen-deuterium exchange (HDX) and arginine scanning to delineate Env-interactive regions of bNAb CAP256-VRC26.03 from donor CAP256 and to model its interaction with V1V2. Finally, we used neutralization screening with UCA and intermediates of V1V2-directed bNAbs to engineer antigens capable of interacting with developmental intermediates. Altogether the structural similarities in antibody recognition along with ontogeny similarities (and differences) in development indicate the V1V2-directed bNAbs to form an ‘extended class,’ which we define as antibodies that do not necessarily share genetic commonalities, but nonetheless display a characteristic mode of antigen interaction. Extended-class immunogens – such as the soluble chimeric trimers we developed here through an ontogeny-based chimera strategy – may provide a general means for eliciting bNAbs against specific sites of Env vulnerability.

RESULTS

Design and crystallization of trimeric scaffolded V1V2s

To obtain a co-crystal structure of HIV-1 Env with a V1V2-directed bNAb from an additional donor, we first screened the V1V2 scaffolds from PDBs 1FD6[28] and 1JO8[29] recognized by bNAbs PG9 and PG16 from donor IAVI 24 and used in their structural analyses[22,25]. Neither bNAbs CH01–04, CAP256-VRC26.01–12, nor PGT141–145 bound these monomeric V1V2 scaffolds. One possible explanation for this lack of recognition might be related to the quaternary dependence of bNAbs CH01–04, CAP256-VRC26.01–12 and PGT141–145, which is more stringent than that of antibody PG9. To accommodate their quaternary dependencies, we designed a library of scaffolded V1V2s based on trimeric scaffolds, expressed secreted version of these scaffolds in a 96-well format[30–32], and assessed PGT145 recognition directly from expressed supernatant (Fig. 1a and Supplementary Fig. 1a, b). PGT145 was chosen for its high quaternary specificity and its broad and potent neutralization, and recognized V1V2 chimeras from two trimeric scaffolds, 1VH8[33] and 4F2K[34].

A 1VH8-scaffolded V1V2 showed strong binding, and three 4F2K-scaffolded V1V2s with differing linker lengths showed moderate binding to PGT145 (Supplementary Fig. 1b, c). We expressed and purified these trimeric scaffolded V1V2s in GnTI- cells and assessed binding by surface plasmon resonance (Biacore), which revealed strong to moderate binding for PG9 (24–500 nM) and weak binding for other V1V2-targeting bNAbs such as CH01 (20–50 μM) (Supplementary Fig. 1d). These scaffolded V1V2s were further optimized by varying HIV-1 strain and linker connections. Optimized 1VH8 trimeric scaffolds with V1V2 from HIV-1 strains CAP256-SU or A244 (which was neutralized well by CH01–04[15]) crystallized with the antigen-binding fragments (Fabs) of bNAbs CH03 or CH04, respectively (Supplementary Fig. 1e).

Structures of scaffolded V1V2 with antibody CH03 and CH04

Diffraction data from co-crystals of Fab CH03 complexed to the scaffolded V1V2 from HIV-1 strain CAP256-SU extended to 3.1 Å, and structure solution by molecular replacement with the previously determined Fab CH04 structure[22] and refinement yielded R_{work}/R_{free} of 21.7%/25.6% (Table 1). Overall, the structure revealed CH03 to insert a protruding CDR H3, bent to recognize strand C of V1V2 in a parallel strand-strand interaction (Fig. 1b).

The scaffolded V1V2 formed a 5-stranded β-barrel, similar to the 5-stranded β-barrel which V1V2 forms in its pre-fusion closed Env conformation[35] as well as the 5-stranded β-barrel formed with antibody 830A[36]. This differs slightly from the 4-stranded Greek Key motif observed in the scaffolded V1V2s bound by PG9[22] and by PG16[25], where the scaffolded V1V2 dimerized at a lattice contact, inhibiting the formation of strand C′. Overall, superposition of V1V2s from scaffolded and Env contexts showed high structural similarity, with a root-mean-square deviation (rmsd) of 1.1 Å for 45 β-strand residues (Fig. 1c).

Analysis of the interface between Fab CH03 and scaffolded V1V2 revealed *N*-glycan interactions to dominate, with 1592 $Å^2$ of the CH03 interactive surface buried by *N*-linked

glycan (82% of the CH03 interface) and with protein of 384 Å^2 (19% of the CH03 interface) (Fig. 2). Electron density was observed for four *N*-linked glycans emanating from residues N130, N156 and N160 of one V1V2 and from residue N139 of a neighboring V1V2 (Fig 2a, b). Over a third of the CH03-glycan interactions occurred with the two protein-proximal N-acetylglucosamine at the base of each *N*-linked glycan; these totaled 9 H-bonded interactions and 741 Å^2 of buried surface (Fig. 2c). With glycan N130, density was observed for four mannose residues, three of which contacted the CH03 heavy chain. Glycan N156 also interacted with the CH03 heavy chain, but no density was observed beyond the protein-proximal N-acetylglucosamines. With glycan N160, density for five mannose residues was observed, of which four were involved with the CH03 heavy chain, and a fifth interacted with the CH03 heavy-light chain interface. Glycan N139 of the neighboring protomer showed clear density for 5 mannose residues and had the most substantial CH03 interactions, comprising 14 H-bonds and 729 Å^2 of buried surface. Four of the N139 mannose residues interacted with the light chain; a fifth was extensively buried at the heavy-light interface with 6 H-bonds and 204 Å^2 of buried surface.

In addition to the glycan interactions, parallel strand-strand interactions occur between a hairpin in the CH03 CDR H3 and strand C of V1V2 (Fig. 2e). Other than the four strand-strand H-bonds, most of the protein-protein interactions were hydrophobic in nature, with Ile100E and Tyr100G of in the CDR H3 stacking over V1V2-strand C backbone and aliphatic portions side chains (Fig. 2f). A salt bridge between Glu30 in the CDR H1 of CH03 and Lys171 of V1V2 provided the only notable electrostatic interaction.

We also obtained co-crystals of Fab CH04 complexed with a 1VH8-scaffolded V1V2 from HIV-1 strain A244. Diffraction data extended to 4.2 Å, and structure solution and refinement yielded R_{work}/R_{free} of 28.5%/32.2% (Table 1). Overall, the CH04–V1V2 structure showed a similar glycan-dominated interface as that of CH03-V1V2, with CH04 also inserting a CDR H3 hairpin to recognize strand C of V1V2 in a parallel strand-strand manner (Supplementary Fig. 2).

Quaternary recognition of glycan in scaffold and Env context

Because CH03 showed extensive contacts with glycan N139 on a neighboring protomer, yet this glycan had not been previously implicated in CH03 interaction, we investigated quaternary scaffold interactions further. Although the V1V2 domain in the trimeric scaffold and trimeric Env align well as individual domains (Fig. 1c), they showed substantially different trimeric arrangement (Fig. 3a, b). Indeed, alignment of a single V1V2 domain from scaffold and Env contexts would position the trimeric counterparts in mostly non-overlapping locations (Fig. 3c).

We observed glycan N139 on the neighboring V1V2 in the trimeric scaffold to be positioned similarly to glycan N160 on the neighboring V1V2 in trimeric Env (Fig. 3d, e). Notably, the protein-proximal N-acetylglucosamines emanating from N139 (scaffold context) and N160 (Env context) were positioned to within 3.5 Å. This suggests that in the trimeric Env context, CH03 would be able to recognize glycan N160 on a neighboring V1V2 in a manner similar to its recognition of N139 in the scaffold context (Fig. 3f, g). In the scaffold context, these CH03 interaction would be symmetric, and lead to three CH03s, each interacting in a similar

way with four glycans. However, in the Env context, the interactions were not compatible with binding of three Fabs, and only a single CH03 Fab would be able to interact with each Env trimer (Fig. 3d, e, inserts).

Modeled interaction between antibody CAP256-VRC26 and V1V2

Similar to the quaternary interaction proposed for CH03 and observed with IAVI 24, V1V2-directed antibodies from both donor CAP256 and donor IAVI 84 also show a single Fab recognizing V1V2 at the trimer apex[19,37]. However, despite substantial effort, we were unable to produce co-crystals for any of the broadly neutralizing antibodies from donors CAP256 or IAVI 84 with V1V2. Modeling of the Fab PGDM1400 structure from donor IAVI 84, which displays a straight extended CDR H3, into a similar orientation as seen in the negative-stained 2D class average EM density of the Fab-Env trimer (PDB 4RQQ[19]) indicated that PGDM1400 CDR H3 may assume a different structure, as the extended CDR H3 observed in the ligand-free Fab crystal structure (stabilized by a strand-strand anti-parallel lattice interaction in the ligand-free Fab context) would otherwise clash. Modeling of the CAP256-VRC26 structures from donor CAP256[16], which display bent CDR H3s more axe-like than hammerhead-like, into negative-stain EM density[16], indicated a possible strand-strand interaction between antibody CDR H3 and V1V2 strand C (Supplementary Fig. 3).

To provide additional experimental constraints, we analyzed rates of hydrogen-deuterium exchange (HDX) from CAP256-VRC26.03 when ligand-free or bound to the BG505 SOSIP.664 Env trimer (Fig. 4a); we used CAP256-VRC26.03 as this was the only structure for which we had structural definition of the entire CDR H3[16]. HDX information from four overlapping peptides pinpointed reduced HDX rates for a 7 residue stretch in the CDR H3; on Env, the HDX information pinpointed a short peptide in the strand C region of V1V2 (Supplementary Fig. 3b). Interestingly, the peptides bearing double tyrosine sulfation appeared to become substantially more protected upon binding, suggesting that the doubly sulfated Fab may form a more stable complex with the epitope than the singly sulfated variant (Fig. 4a). We also used an arginine scanning approach[22] to map the CAP256-VRC26.03 paratope (Supplementary Fig. 3c); arginine scanning disrupts interactions through steric interference allowing for a more robust delineation of contacts than alanine scanning. This paratope mapping implicated the CDR H3 as the primary means of Env contact (Fig. 4b).

We combined constraints from Fab-Env negative stain, from HDX, and from arginine scanning to model the CAP256-VRC26 interaction with V1V2 in the context of the BG505 SOSIP.664 trimer (Fig. 4c). We used the CAP256-VRC26.09 antibody for this modeling, as this antibody showed high affinity for the SOSIP trimer, and the negative stain EM map with the SOSIP.664 trimer was obtained with this antibody. This model suggested a strand-strand interaction between the seven residues of the CDR H3 implicated by HDX and strand C of V1V2, with the strand-strand interaction occurring in an antiparallel manner (Fig. 4d). Trp side chains could stack against the strand C backbone, with electrostatic interactions facilitated by two sulfated Tyr and two Asp residues; because of the greater distance between strand-strand interactions and other CDR regions, interactions with *N*-linked glycan

would be expected to occur with the mannose tips, more than the N-acetylglucosamine stalks.

A common mode of V1V2 recognition by antibody

The co-crystal structures of V1V2 with CH03 and CH04 from donor CH0219 and the modeled interaction of V1V2 with CAP256-VRC26 from donor CAP256 provide atomic-level insight into V1V2-directed bNAbs from two donors, and we integrated this information into the previously published characterization of V1V2 with PG9 and PG16 from donor IAVI 24 (Fig. 4e, f, g). In all three donors, a strand-strand interaction was implicated between an extended CDR H3 on antibody and the C strand of V1V2. With bNAbs from donors CH0219 and IAVI 24, the interaction involved parallel hydrogen-bonding, whereas in donor CAP256, the interaction involved anti-parallel hydrogen bonding. In terms of CDR H3 side-chain interactions, bNAbs from CH0219 utilized primarily hydrophobic interactions, those from IAVI 24 primarily electrostatic interactions, and those from CAP256 a mixture of hydrophobic and electrostatic interactions.

bNAbs from all three donors bound a single Fab per trimer (Supplemental Fig. 4). With bNAbs from CH0219, the crystal structure data indicated this stoichiometry and known quaternary-specificity to derive from recognition of *N*-linked glycan that emanate from neighboring V1V2s. With bNAbs from IAVI 84, the published structural data[22,24,25] also indicated quaternary-recognition of N160. With bNAbs from CAP256, modeling indicated protein-pr otein interactions to occur mostly within a single V1V2 protomer, however the model implicated electrostatic interactions with each of the neighboring protomers suggesting the strong quaternary specificity of the CAP256-VRC26 bNAb lineage to derive from quaternary recognition of glycan and protein (Supplementary Fig. 4). Minor differences in the precise engagement of the glycans between IAVI 24 and CH0219 may result in varying tolerances to different glycoforms present on the HIV-1 Env, as has been seen with V3 N332-directed lineages.

We observed some CDR H3 features, such as an enrichment of Asp and Tyr residues in the CDR H3s of all of the V1V2-directed antibodies. These appeared to play diverse roles. For example, a D-gene encoded YYD motif occurred in the CDR H3s of both PG9 and CAP256-VRC26.09, and in both antibodies both tyrosines are sulfated. However, in PG9, the sulfated tyrosines interact with strand C sidechains; in CAP256-VRC26.09 they appeared to interact with a neighboring protomer (Supplementary Fig. 4). The conservation of the YYD motif despite its different roles may relate to the preferred anionic character of the paratope[16,22], with the YYD motif incorporating potentially three negative charges, if the tyrosines were sulfated, as they are in PG9 and CAP256-VRC26.09. Also, the preferred use of specific D and J genes, such as D3-3 which contains the YYD motif, is associated with long CDR H3s[38,39]. Overall, while there were substantial differences between the V1V2-directed antibodies, the similarities that we observed in strand-strand protein recognition and quaternary recognition of glycan suggested a common mode of V1V2 recognition by bNAbs from donors CH0219, CAP256 and IAVI 24.

Ontogenies of V1V2-directed antibodies

The finding of a common mode of V1V2 recognition, shared by bNAbs from multiple donors in the absence of a common genetic origin, suggested an “extended class”, which we define as antibodies lacking specific genetic commonalities, but nonetheless utilizing a characteristic mode of antigen interaction. We further investigated the development of these V1V2-directed bNAbs to see if similarities in molecular recognition might extend to similarities in B cell ontogeny. In particular, we sought to understand how the structural characteristics of the V1V2-directed bNAbs such as a long, anionic CDR H3s – needed to penetrate the glycan shield and interact with the exposed cationic strand C in V1V2 – were created by B cell processes of recombination and somatic hypermutation and whether these processes were similar from donor to donor.

To enhance the lineage information of V1V2-directed bNAbs from donor CH0219, we used both paired and unpaired heavy-light chain sequencing to obtain sequences of additional lineage members. The new sequences were most similar to antibody CH01 (Supplementary Fig 5). The earliest well-defined maximum-likelihood intermediate (I) had 0.3% nucleotide-level somatic hypermutation, with a CDR H3 of 24 amino acids and a net –4 charge (Fig. 5a, top). Analysis of glycan-interacting residues indicated 71% to be present by intermediate I (Supplementary Fig. 5). Altogether the lineage information indicated the long CDR H3 to derive from recombination, with residues involved in glycan recognition to be partially present in the initial recombinant and to evolve upon maturation.

The CAP256-VRC26 bNAb lineage, previously substantially defined by longitudinal sampling[16], also indicated a long CDR H3 to be the result of recombination (Fig 5a, second row). For the PG9-bNAb lineage from IAVI 24, which has been defined with only two members, PG9 and PG16, we also carried out maximum likelihood analysis. The earliest intermediate (I) had 4.4% nucleotide-level somatic hypermutation, with a CDR H3 of 28 amino acids and a net –3 charge (Fig. 5a, third row). Analysis of glycan-interacting residues indicated 43% to be present by the earliest intermediate (Supplementary Fig. 6).

We also analyzed the PGT145-bNAb lineage from donor IAVI 84 (Supplementary Fig. 5). Although the likely conformational change in CDR H3 between free and bound conformations made modeling difficult, many features of PGT145-lineage antibodies were in common with V1V2-directed bNAbs from CH0219, CAP256 and IAVI 24, suggesting a similar mode of recognition. To the PGT141–145 and PGDM1400–1412 series of bNAbs[18,19], we used “intra-donor” phylogenetic analysis to add sequences of antibodies from previously published next-generation sequencing of donor IAVI 84-B cell transcripts[40]. The earliest well-defined maximum-likelihood intermediate (I) had 6.1% nucleotide-level somatic hypermutation, with a CDR H3 of 32 amino acids and a net –5 charge (Fig. 5a, bottom row).

To determine which class features might provide roadblocks for development, we analyzed the prevalence of antibodies with long anionic CDR H3s and high SHM from normal donors[41] and compared these with the four V1V2-directed bNAb lineages (Figs. 5b). In the CDR H3 length and charge distributions, bNAbs from donors CH0219 and IAVI 24 were at the edge of normal, whereas bNAbs from donors CAP256 and IAVI 84 showed more

extreme values. In the SHM histogram, all antibodies were at the upper range of normal, except for PGDM1400 which was substantially higher. Altogether the lineage information from all four donors indicated V1V2-directed bNAbs to start with a long anionic CDR H3 at the recombination stage, which in a number of donors appeared to have partial recognition of *N*-linked glycan, which evolved during maturation.

HIV-1 strains recognized by developmental intermediates

We previously found reversion of V1V2-directed bNAbs to UCA (or to V-gene reverted approximations) to substantially restrict neutralization potency[15,16,20]. However, prior neutralization analyses were not performed with antibodies from all four donors against the same panel of viruses. We therefore assessed the neutralization of mature, intermediate, and UCA versions of V1V2-directed bNAbs from donors CH0219, CAP256, IAVI 24 and IAVI 84 on a panel of roughly 200 HIV-1 isolates (Fig. 6a).

To define a common set of HIV-1 interactive isolates, we used a frequentist analysis to determine the likelihood that a particular strain might be recognized by a reverted bNAb. This analysis identified nine strains recognized by many of the reverted bNAbs (Fig. 6a, inset). These strains were ~10–20 times more likely than the median of the panel to be neutralized by the earliest neutralizer from each lineage (Supplementary Table 1). The clade B strain WITO was recognized commonly, and the clade C ZM233 strain, which we had earlier shown to be recognized by reverted versions of PG9 and CH01[15,42], was also commonly recognized. Direct analysis of strand C sequences for these strains showed some differences (e.g. ZM233 does not contain a glycan at residue N156 or N173, differentiating it from others in the group) and many commonalities, such as the presence of glycan N160 and a cationic N-terminal half of strand C (Supplementary Fig. 7). However, we did not observe a specific distinguishing sequence characteristic separating strains neutralized by reverted bNAbs from those not neutralized. In general, the V1V2 epitope of these bNAbs involves glycan, is quaternary, and is known to adopt multiple conformations. For instance strand C, in a monomeric context, can adopt multiple conformations[43], and the V1V2 trimeric interface on virions is conformationally dynamic[44]. Thus, the mechanism by which specific strains interact more strongly with reverted V1V2-directed bNAbs is likely both complex and multi-factorial. That said, we did observe that neutralization by reverted bNAbs correlated with the neutralization potency of the mature bNAbs – and thus one distinguishing characteristic was the high susceptibility of the select strains to being neutralized, not only by the reverted bNAbs, but also by the mature ones (this susceptibility was specific to V1V2 directed antibodies, as these strains were not generally sensitive to neutralization). In general, the direct experimental assessment of strains capable of interacting with UCA or reverted bNAbs through a neutralization-based screen (such as that shown in Fig. 6) identifies strains from which to engineer antigens capable of binding early members of V1V2-directed bNAb lineages.

Design and validation of ontogeny-specific chimeric trimers

While the trimeric V1V2 scaffolds we designed and optimized for crystallization might have utility as bNAb-specific probes, their altered oligomeric presentation of V1V2 (Fig. 3) along with neo-epitopes from both scaffold and exposed regions of V1V2 suggest that they might

be less than optimal as vaccine immunogens. This contrasts with the correct oligomeric presentation of V1V2 in SOSIP.664 and DS-SOSIP.664 contexts[27,32,35] and the ability of SOSIP.664 to elicit high titers of Tier 2 autologous neutralizing titers[45]. We therefore chose to use soluble DS-SOSIP.664 trimers as a platform to obtain ontogeny-specific antigens.

One issue with DS-SOSIP.664 is a difficulty in expressing strains other than BG505. To address this, we used our understanding of the V1V2 structure in the context of the pre-fusion closed HIV-1 Env trimer to construct V1V2 chimeras, which swapped desired strains of HIV-1 Env into the V1V2 region of BG505 DS-SOSIP.664 trimers (Fig. 6b). We succeeded in expressing chimeric DS-SOSIP trimers with V1V2s from all nine of the special strains we identified by lineage reversion and HIV-1 neutralization (Fig. 6a and Supplementary Fig 7). We assessed recognition by mature and reverted bNAbs of all nine V1V2 chimeras (Table 2 and Supplementary Fig 7). Binding of the reverted bNAbs to each of the nine V1V2 chimera correlated with their neutralization of the same nine HIV-1 strains (spearman ρ=0.67; $p< 0.0001$).

Discussion

The path to creating a set of antigens capable of interacting with germline-reverted versions of V1V2-directed bNAbs involved an information flow from a supersite of vulnerability, through the definition of an extended class and the identification of specific strains capable of being neutralized by germline-reverted class members, to the engineering and antigenic validation of V1V2-chimeric trimers (Fig. 7a). We note that for V1V2-directed bNAbs of the PG9-extended class (named for the first described antibody of the extended class), ontogeny analysis indicates three likely roadblocks to development. The first roadbock, recombination to produce a long anionic CDR H3 capable of piercing the glycan shield to backbone-hydrogen bond to strand C of V1V2 (Fig. 7b, left column) is likely a substantial impediment in the development of V1V2-directed bNAbs due in part to their rarity as a result of tolerance deletion from the naïve B cell repertoire[46]. Next-generation sequencing and CDR modeling, however, have suggested suitable CDR H3s to be present in the naïve human repertoire[38]. Our ontogeny-based immunogens are not expected to influence recombination frequencies, and thus likely has little impact on this B cell stage.

A second roadblock –B cell priming involves an interaction between the initial recombinant on a naïve B cell and HIV-1 Env of sufficient strength to prime the B cell lineage (Fig. 7b, middle column)– appears to be a substantial impediment. With all four bNAb V1V2-directed lineages analyzed here, the UCA or gHgL could neutralize less than 5% of HIV-1 isolates (Fig. 6a), suggesting that less than 1 in 20 HIV-1 strains can prime the naïve B cell of a potential V1V2-directed bNAb lineage. In the case of the CAP256-VRC26 lineage, sequencing of co-evolving antibody and virus indicates that successful priming of this lineage only occurred because the sequence of a superinfecting virus (CAP256-SU) fortuitously happened to interact with the lineage ancestor[16]. Immunization with soluble trimers with V1V2 chimeras of the special strains capable of interacting with initial recombinant, however, should substantially increase the likelihood successful priming of the lineage. Individually, the enhancement in likelihood of interaction is ~10–20 times higher for each of the nine special strains we identify here (Fig. 6a); cocktails of these strains are

expected to increase the likelihood even more, with a nine-strain cocktail of all of the special strains enhancing the probability of interaction by ~40-fold (Supplementary Table 1b).

A third roadblock, somatic development through which the primed B cell matures from having weak interactions with few HIV-1 strains to having strong interactions with a majority of HIV-1 strains (Fig. 7b, right column) is likely to be a substantial impediment as well, with each of the bNAb lineages showing substantial maturation prior to the emergence of mature antibodies capable of neutralizing a majority of HIV-1 strains (Fig. 5a), and longitudinal data from the CAP256-VRC26 lineage indicating years of maturation[16]. It should be possible to quantify the maturation induced by our V1V2 chimeric trimers in mouse systems, comprising knockins of germline or early intermediates of the V1V2-directed antibodies, as has been done with knockins of germline or early intermediates for other bNAbs, such as of the VRC01 class[12,13].

The ability of ontogeny-specific vaccine immunogens to affect specific stages of B cell development, such as shown for the V1V2-extended class (Fig. 7c), may be generally applicable to extended bNAb classes for which steps of UCA priming or of somatic maturation represent roadblocks to antibody development[3,5,7,12,13]. Currently the only classes of broad HIV-1-neutralizing antibodies that have been observed involves the VH1-2-derived VRC01 class and the VH1-46-derived 8ANC131 class, both of which target the CD4 supersite[8,10,11]. For the VRC01 class, a severe roadblock to development involves recognition by the UCA of HIV-1 Env[5,7,10,47]. Nevertheless, neutralization by VRC01-class reverted intermediates identified strains with characteristics known to enhance UCA recognition, such as the deletion of glycan 276 in gp120 loop D[5,10,48] (Supplementary Fig. 8). It will be exciting to see if other classes or extended classes of bNAbs targeting HIV-1 will be identified – and if the reversion-based neutralization strategy for ontogeny-specific vaccine design described here will be useful in their “re”-elicitation.

Online Methods

Protein expression, purification

V1V2 scaffolds, codon optimized for mammalian expression, were designed with an artificial N-terminal secretion signal and a C-terminal HRV3C recognition site followed by an 8-His tag. The genes were cloned into the XbaI/BamHI sites of the mammalian expression vector pVRC8400, and transiently transfected into HEK 293S GnTI- cells, which were used due to a requirement for a Man5GlcNAc2. Scaffolds were purified from the media using Ni-NTA resin (Qiagen), and the eluted proteins were digested with HRV3C(Novagen) before passage over a 16/60 S200 size exclusion column.

CH03 and CH04 heavy and light chains were codon optimized for mammalian expression with an HRV3C recognition site (GLEVLFQGP) inserted after Lys 235 of the heavy chain. The sequences were cloned into pVRC8400 and transiently co-transfected into Expi293 cells as previously described[22]. IgG was purified from the supernatant after 5 days using Protein A agarose (Pierce) and dialyzed into PBS.

To form homogeneous complexes, 3 mg of purified CH03 or CH04 HRV3C IgG was bound to 750 ml Protein A Plus agarose (Pierce) in a disposable 10 ml column. To this resin was added 10 mg of purified V1V2 scaffold. Unbound scaffold was washed through the column with 5 column volumes of PBS. The column was capped and 20 μl of HRV3C protease at 2U μl^{-1} was added to the resin along with 1 ml of PBS. After 2 h at room temperature, the resin was drained, the eluate collected and passed over a 16/60 S200 column in buffer containing 5 mM HEPES 7.5, 150 mM NaCl, and 0.02% NaN_3. Fractions corresponding to the CH03–V1V2 complex were pooled and concentrated to 5 mg ml^{-1}.

Crystallization screening

CH03-1VH8-V1V2$_{SU}$ complex and CH04-1VH8-V1V2$_{A244}$ were screened for crystallization using 572 conditions from Hampton, Wizard and Precipitant Synergy[49] screens using a Cartesian Honeybee crystallization robot as described previously[22] and a mosquito robot using 0.1 μl of reservoir solution and 0.1 μl of protein solution. Crystals for CH03-1VH8-V1V2$_{CAP256SU}$ suitable for structural determination were obtained in 10% PEG 8000, 0.1 M CaCl, 20% MPD, 0.1M Na Acetate pH 5.5. Crystals were cryoprotected in a solution supplemented with 15% Sucrose and flash-frozen in liquid nitrogen. Crystals for CH04-1VH8-V1V2$_{A244}$ suitable for structural determination were obtained in 54% Isopropanol and 0.1 M Tris pH 8.5. Crystals were cryoprotected in a solution supplemented with 25% (2R,3R)-butane-2,3-diol and covered in paratone oil before being 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

Diffraction data were processed with the HKL2000 suite[50]. A molecular replacement solution for the data set consisting of three CH03 Fab molecules per asymmetric unit with the 1VH8 trimer was obtained using Phenix. The CH04 complex contained 2 Fabs and 2 monomers of 1VH8, the trimer conformation was maintained through crystal symmetry. Model building was carried out using COOT, and refinement was performed with Phenix. Final data collection and refinement statistics are presented in Table 1. The Ramachandran plot as determined by Molprobity shows 91% of all residues in favored regions and 97.2% of all residues in allowed regions for the CH04 structure and 92% of all residues in favored regions and 98.4% of all residues in allowed regions for the CH04 structure.

Antigenic analysis of BG505 SOSIP.664.DS.V1V2 by MSD-ECLIA

Standard 96-well bare MULTI-ARRAY Meso Scale Discovery (MSD) Plates (MSD, cat #L15XA-3) were coated with a panel of HIV neutralizing (VRC01[6], b12[51], PGT121[18], PGT128[18], 2G12[52], PGT145[18], CAP256-VRC26.25[53], 35O22[54] and 8ANC195[55]), non-neutralizing monoclonal antibodies (F105[56], 17b[58] and 447-52D[59]) and non-cognate antibodies (anti-influenza antibody CR9114[60] and anti-RSV antibody D25[30]) in duplicates (30 μL/well) at a concentration of 4 μg/mL, diluted in 1X PBS and the plates were incubated overnight at 4°C. The following day, plates were washed (wash buffer: 0.05% Tween-20 +1 X PBS) and blocked with 150 μL of blocking buffer [5% [W/V] MSD Blocker A (MSD, Cat # R93BA-4)] and incubated for 1 hr on a vibrational shaker (Heidolph TITRAMAX 100;CAT # P/N: 544-11200-00) at 650 rpm. All the incubations were performed at room

temperature, except the coating step. During the incubation, BG505 SOSIP V1V2 chimeric trimers were titrated down in a serial 2X dilutions starting at 5 μg/mL concentration of the trimer in assay diluent (1% [W/V] MSD blocker A + 0.05 % Tween-20). After the incubation with blocking buffer was complete, the plates were washed and the diluted trimer was transferred (25 μl/well) to the MSD plates and incubated for 2 hrs on the vibrational shaker at 650 rpm. After the 2 hr incubation with trimer, the plates were washed again and 2G12 antibody labeled with MSD Sulfotag (MSD; Cat #R91AN-1) at a conjugation ratio of 1:15 [2G12: Sulfotag]), which was diluted in assay diluent at 2 ug/mL, was added to the plates (25 μL/well) and incubated for 1 hr on the vibrational shaker at 650 rpm. The plates were washed and read using the 1X read buffer (MSD Read Buffer T (4x); Cat# R92TC-2) on MSD Sector Imager 2400.

Antigenic analysis of BG505 SOSIP.664.DS.V1V2 by ELISA

96-well nickel coated plates (Thermo Fisher Scientific) were coated for 1 h at 100 μl/well with his-tagged BG505 SOSIP.664.DS.V1V2 constructs at 1 μg/ml, diluted in 1× PBS. The plates were then blocked at room temperature for 1 h using 200 μl/well of 5% skim milk in 1× PBS. After the plates were washed (wash buffer, 0.05% Tween 20 plus 1× PBS), they were incubated for 1 h at room temperature with 5-fold serially diluted V1V2-directed antibodies at concentrations ranging from 20 to μg/ml to 1 ng/ml in 1 × PBS plus 0.05% Tween 20. The plates were washed and then incubated for another hour with horseradish peroxidase (HRP)-conjugated anti-human IgG (1:5000) diluted in 0.2% Tween 20 plus 1× PBS. After a final wash, the plates were then developed using TMB peroxidase substrate for 20 min. The reaction was stopped with 140 mM HCl, and then the absorbance was measured at 450 nm. All incubations were at 100 μl/well at room temperature, except where noted otherwise. Experiments were run in triplicate.

Neutralization Panels

Single round of replication Env-pseudoviruses were prepared, titered, and used to infect TZM-bl target cells as described previously[63,64]. The data were calculated as a reduction in luminescence units compared with control wells, and neutralization curves were fit by nonlinear regression using a 5-parameter Hill slope equation. Data are reported as 50% inhibitory concentration (IC_{50}) in micrograms per microliter. Neutralization was assayed using a previously described panel[6] of 200 geographically and genetically diverse Env-pseudoviruses representing the major subtypes and circulating recombinant forms.

Negative-stain electron microscopy

Negative-stain electron microscopy samples were diluted to about 0.03 mg/ml, adsorbed to a freshly glow-discharged carbon-film grid for 15s, and stained with 0.7% uranyl formate. Images were collected semi-automatically using SerialEM[65] on a FEI Tecnai T20 with a 2k x 2k Eagle CCD camera at a pixel size of 0.22 nm/px. Particles were picked automatically and reference-free 2D classification was performed in EMAN2[66].

Hydrogen Deuterium Exchange (HDX)

To identify Env-interactive residues within CAP256-VRC26.03, rates of hydrogen-deuterium exchange for CAP256-VRC26.03 Fab alone and in the presence of BG505 SOSIP.664 (2:1 molar ratio, Fab:SOSIP - a ratio for which 50% of CAP256-VRC26.03 Fab and 100% of SOSIP should be bound) were measured. Deuteration reactions were initiated by diluting 10 uL of 0.56 mg/mL CAP256-VRC26.03 Fab (+/− SOSIP) into 90 uL of phosphate buffered saline containing 85% D_2O. Deuterations were carried out for 10 sec, 12 sec, and 60 sec prior to the addition of 90 uL of quench solution (500 mM Tris-2-carboxyethyl phosphine (TCEP), 3M urea, 0.2% formic acid, with a final pH 2.5), 10 uL pepsin (2 mg/mL in 100 mM phosphate pH 4.0) and digestion on ice for 5 minutes. Pepsin-digested HDX samples were plunge frozen in liquid nitrogen and stored at −80C. Unliganded and SOSIP-bound VRC26.03 HDX samples were prepared side-by-side in matched buffers. For LC-MS analysis, samples were thawed on ice for 5 minutes and manually injected onto a Waters BEH 1.7 μm 1.2 x 5 mm trap column (Waters) flowing 0.1% TFA at 200 μL/min. After 3 minutes of washing the peptides were resolved over a Hypersil 1 x 50 mm 2.1 μm C18 column (Thermo Scientific) using a gradient of 15 to 40% B in 8 minutes (A: 0.05% TFA 5% ACN; B: 0.05% TFA 80% ACN). Eluted peptides were analyzed in duplicate with a Waters Synapt Q-TOF mass spectrometer. Peptide identification and exchange analysis were as described previously[67], with the exception that percent deuteration was calculated using the number of exchangeable backbone amides to estimate theoretical 100% deuteration for each peptide.

in-silico Arginine scanning

Step one of our *in-silico* arginine scanning method involves mutating each residue in the CDR region to arginine. Any endogenous arginine or lysine residues were left unmutated but still used in all subsequent analysis. To carry out all of our in silico mutations, we used the side chain modeling program *Scap*[68] with default settings. Step two of our method scores and determines the solvent accessibility of each predicted arginine residue. To assess the feasibility of adopting an arginine at a particular position in a CDR region, each mutated residue was scored independently using a modified *RAPDF* distance-dependent statistical potential[41]. The scores for each residue were computed as follows: $\Delta E_i = RAPDF_i(ARG) - RAPDF_i(NAT)$ where the subscript *i* denotes the residue position and *ARG* and *NAT* denote the score of the mutated and native residue respectively. We used the *Surface*[69] program to determine the solvent accessible surface area (SASA) of each residue. The SASA was normalized using a reference SASA for the same residue determined from a linear peptide composed of three residues in extended conformation. The first and last residue of the linear peptide is glycine. The normalized SASA was computed as follows: $Acc_i = SASA_i(S)/SASA_i$ where *S* denotes the sidechain in the protein environment and *L* denotes the side chain in the linear peptide environment. The normalization factor allows for an unbiased comparison of solvent accessibility for residues of different amino acid types (and sizes). A weighting factor (set to 1.0) was used to convert the normalized SASA value into a score. The final step of our method involves a stochastic search for the optimal set of arginine residues based on the RAPDF and solvent accessibility scores determined in step two. Each cycle of the search begins by randomly selecting a seed residue from the set of CDR residues. All residues within a 5.0 Å radius of the seed residue were removed (‘flagged’) from the set of

selectable residues in the current cycle. We next determined the closest residue greater than 5.0 Å (measured from the α-carbon atom) away from the current seed residue. This residue was now set as the seed residue and all residues within a 5.0 Å radius were removed from the set of selectable residues. The process is repeated until all residues have been selected. At the end of each cycle, the total score for the final set of seed residues was determined from the look-up tables generated in step two. The total score was computed as follows: $F=\Sigma_{i=1,N}(\Delta E_i - Acc_i)$ where ΔE_i is the RAPDF –based score and Acc_i is the weighted-normalized SASA score. The stochastic search will continue until the lowest F score no longer changes or the user-specified maximum number of cycles (set to 1000) has been reached. The set of residues with the lowest F score was selected as the optimal set of residues to mutate to arginine.

Phylogenetic analysis involving donors CH0219, CAP256, IAVI24 and IAVI84 and CAP256

Donor CH0219—All 454 and Illumina NGS heavy chain sequences were processed using an in-house version of ***IgBlast*** [ref PMID: 23671333] and then filtered based on their heavy chain V and J germline gene assignments. All sequences with the IgVH 3–20 and IgJH 2 germline gene assignments were retained for further processing. To enrich the pool of transcripts with long CDR H3s, we filtered out any sequences with CDR H3 lengths (Kabat) less than 24 amino acids. Duplicated sequences were removed and problematic sequences were edited to bring the transcript into the correct translational frame. A multiple sequence alignment for the final set of sequences was generated using ***CLUSTAL0***[70] with default settings followed by manual inspection. A maximum-likelihood (ML) tree was constructed using the program ***dnaml*** which is part of the PHYLIP package v3.69 (http://evolution.genetics.washington.edu/phylip.html). The 'S' option was set to No to provide a more thorough optimization. The inferred intermediates were derived from the ML tree. Nucleotide sequences for CH01–CH04 were downloaded from GenBank (accession # JQ267523.1-JQ267526.1).

Donor CAP256—The ML tree was obtained from the study of Doria-Rose et al. [16] The nucleotide sequence for the CAP256-VRC26.UCA was downloaded from GenBank (accession #: KJ134860.1).

Donor IAVI24—nucleotide sequences for PG9 (accession #: GU272045.1) and PG16 (accession #: GU272043.1) were aligned and the alignment used to generate an ML tree using the program ***dnaml*** using the same procedure as was done for donor CH0219. The inferred intermediate was derived from the ML tree.

Donor IAVI84—454 NGS sequences were downloaded from the SRA (accession # SRP018335) and processed in the same manner as was done for donor CH0219. Sequences where then filtered according to their V germline gene assignments. All sequences with the IgVH1-8 germline gene assignment were retained for further processing using intra-donor phylogenetic analysis. The major objective of intra-donor phylogenetic analysis is to bracket all phylogenetically similar sequences on a Neighbor-Joining (NJ) tree using known neutralizing antibody sequences derived from the same donor. For the analysis here, we used all possible pairs of neutralizing antibody sequences derived from this donor

PGDM1400-1412 (accession #: KP006370–KP006382) and PGT141–145 (accession #: JN201906.1-JN201910.1). Intra-donor phylogenetic analysis works in similar fashion to the cross-donor phylogenetic analysis described previously[11]. Briefly, the method begins by randomly shuffling all the sequences in a data set to remove any potential bias in the order of the sequences and improves the convergence of the method. After sequence shuffling, the data set were split into FASTA files each containing up to 5000 sequences. A pair of neutralizing antibody sequences along with the germline VH gene sequence was added to each FASTA file. The germline gene sequence was used as the outgroup in the NJ tree. A multiple sequence alignment for each FASTA file was generated using ***CLUSTAL0***[70] and each alignment was then converted into a distance matrix using the ***dnadist*** program (with default settings). From this, a NJ tree was constructed using the ***neighbor*** program (with default settings). Both ***dnadist*** and ***neighbor*** are part of the PHYLIP package. Donor sequences were extracted from each NJ tree using a pair of neutralizing antibody sequences derived from the same donor. All donor sequences contained in the minimal-spanning tree containing the pair of neutralizing sequences were extracted from the NJ tree and split into FASTA files containing no more than 5000 sequences and the above process repeated five times. The remaining sequences were filtered by their germline J gene retaining all sequences with the IgJH 6 assignment. We filtered out any transcripts with CDR H3 lengths (Kabat) less than 31 amino acids and greater than 34 amino acids. Duplicate sequences were removed and problematic sequences were edited to bring the transcript into the correct translational frame. The remaining sequences were aligned and then used to generate an ML tree using the program ***dnaml*** using the same procedure as was done for donor CH0219. The intermediates were derived from the ML tree.

Modeling the CAP256-VRC26.9 the HIV-1 Env trimer complex

A combination of molecular dynamics (MD) and molecular dynamics flexible-fitting (MDFF) simulations were used to model the complex consisting of the prefusion HIV-1 Env trimer and CAP256-VRC26.09 antibody. To begin, any loop regions on the HIV-1 Env prefusion trimer were modeled in using the program Loopy[71]. We then added Mannose 5 to each sequon position. A model for the CAP256-VRC26.09 antibody was obtained by threading the amino acid sequence of CAP256-VRC26.09 onto the CAP256.VRC26.03 antibody structure (PDB ID: 4OD1) using the program Nest[69]. Models for the trimer and antibody were docked as rigid-bodies into the cryo-EM map (EMD #: 5856) using the ***Colores*** program within the SITUS package[72]. The complex was then solvated in a 15Å water box and neutralized by the addition of NaCl at a concentration of 150 mM.

After rigid-body docking into the cryo-EM map, we applied a combination of MDFFand MD to the complex followed by MD alone to relax the complex. The simulation was performed using NAMD[73], with the CHARMM36 force field, including CMAP corrections for the protein. TIP3P water parameterization was used to model the solvent. Periodic electrostatic interactions were computed using particle-mesh Ewald (PME) summation with a grid spacing smaller than 1Å. Constant temperature was imposed by using Langevin dynamics with a damping coefficient of 1.0 ps. A constant pressure of 1 atmosphere was maintained with Langevin piston dynamics, 200 fs decay period and 50 fs time constant. During equilibration, the trimer’s secondary structure was restrained with harmonic

restraints applied to the backbone's torsional angles and interacting strands (strand C on the trimer and the strand on CDR H3). The system was minimized by 5000 conjugate gradient steps and then equilibrated by using a linear temperature gradient, which heated up the system from 0 to 310 K in 5 ns. An additional 5 ns of simulation was carried out followed by 5 ns MDFF. The proteins secondary structure, the chirality and the interacting strains were restrained. The system was relaxed during 20 ns and then subjected to a second 5 ns MDFF run, where the interacting strains were not restrained. Finally, unrestrained molecular dynamics were performed up to 50 ns. The length of all bonds involving hydrogen atoms was constrained using the RATTLE algorithm, thus allowing a time step of 2 fs.

Figures

Structure figures were prepared using PyMOL ((http://www.pymol.org/). Figure 6a, was generated using GraphPad Prism (http://www.graphpad.com/scientific-software/prism/). EM structure images were generated using Chimera (http://www.cgl.ucsf.edu/chimera/). Phylogenetic trees were generated using Dendroscope (http://ab.inf.uni-tuebingen.de/software/dendroscope).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank members of the Structural Biology Section and Structural Bioinformatics Core, Vaccine Research Center for discussions and comments on the manuscript, and the Weill Cornell Medical College, The Scripps Research Institute (TSRI), Academic Medical Center and the HIV Vaccine Research and Design team for their contributions to the design and validation of near-native mimicry for soluble BG505 SOSIP.664 trimers. We thank J. Baalwa, D. Ellenberger, F. Gao, B. Hahn, K. Hong, J. Kim, F. McCutchan, D. Montefiori, L. Morris, J. Overbaugh, E. Sanders-Buell, G. Shaw, R. Swanstrom, M. Thomson, S. Tovanabutra, C. Williamson, and L. Zhang for contributing the HIV-1-Env plasmids used in our neutralization panel. We thank Rogier Sanders for providing the PGDM1400-12 sequences and the International AIDS Vaccine Initiative (IAVI) for PG9, PG16 and PGT141–145;. Support for this work was provided by the Intramural Research Program of the Vaccine Research Center, US National Institute of Allergy and Infectious Diseases (NIAID) (to A.B.M., J.R.M. and P.D.K.); the Division of AIDS NIAID, NIH (1U01-AI116086-01 to P.L.M., L.M., J.R.M. and P.D.K.; R21-AI112389 to K.K.L;); the Bill and Melinda Gates Foundation Collaboration for AIDS Vaccine Discovery (OPP1033102 to K.K.L.); IAVI; and the Center for HIV/AIDS Vaccine Immunology-Immunogen Discovery grant (CHAVI-ID; UM1 AI100645 to M.B. and B.F.H.). This project was funded in part with Federal funds to U.B. from the Frederick National Laboratory for Cancer Research, NIH, 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-31-109-Eng-38. Modeling and molecular dynamics were carried out using the NIH's Biowulf computing cluster.

References

1. Hraber P, et al. Prevalence of broadly neutralizing antibody responses during chronic HIV-1 infection. AIDS. 2014; 28:163–9. [PubMed: 24361678]
2. Burton DR, et al. HIV vaccine design and the neutralizing antibody problem. Nat Immunol. 2004; 5:233–6. [PubMed: 14985706]
3. Haynes BF, Kelsoe G, Harrison SC, Kepler TB. B-cell-lineage immunogen design in vaccine development with HIV-1 as a case study. Nature biotechnology. 2012; 30:423–33.
4. Kwong PD, Mascola JR. Human antibodies that neutralize HIV-1: identification, structures, and B cell ontogenies. Immunity. 2012; 37:412–25. [PubMed: 22999947]

5. Jardine J, et al. Rational HIV immunogen design to target specific germline B cell receptors. Science. 2013; 340:711–6. [PubMed: 23539181]
6. Wu X, et al. Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. Science. 2010; 329:856–61. [PubMed: 20616233]
7. Zhou T, et al. Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. Science. 2010; 329:811–817. [PubMed: 20616231]
8. Scheid JF, et al. Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding. Science. 2011; 333:1633–7. [PubMed: 21764753]
9. Wu X, et al. Focused evolution of HIV-1 neutralizing antibodies revealed by structures and deep sequencing. Science. 2011; 333:1593–602. [PubMed: 21835983]
10. Zhou T, et al. Multidonor analysis reveals structural elements, genetic determinants, and maturation pathway for HIV-1 neutralization by VRC01-class antibodies. Immunity. 2013; 39:245–58. [PubMed: 23911655]
11. Zhou T, et al. Structural Repertoire of HIV-1-Neutralizing Antibodies Targeting the CD4 Supersite in 14 Donors. Cell. 2015; 161:1280–92. [PubMed: 26004070]
12. Dosenovic P, et al. Immunization for HIV-1 Broadly Neutralizing Antibodies in Human Ig Knockin Mice. Cell. 2015; 161:1505–15. [PubMed: 26091035]
13. Jardine JG, et al. HIV-1 VACCINES. Priming a broadly neutralizing antibody response to HIV-1 using a germline-targeting immunogen. Science. 2015; 349:156–61. [PubMed: 26089355]
14. Kong L, et al. Supersite of immune vulnerability on the glycosylated face of HIV-1 envelope glycoprotein gp120. Nat Struct Mol Biol. 2013; 20:796–803. [PubMed: 23708606]
15. Bonsignori M, et al. Analysis of a clonal lineage of HIV-1 envelope V2/V3 conformational epitope-specific broadly neutralizing antibodies and their inferred unmutated common ancestors. J Virol. 2011; 85:9998–10009. [PubMed: 21795340]
16. Doria-Rose NA, et al. Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies. Nature. 2014; 509:55–62. [PubMed: 24590074]
17. Walker LM, et al. Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. Science. 2009; 326:285–9. [PubMed: 19729618]
18. Walker LM, et al. Broad neutralization coverage of HIV by multiple highly potent antibodies. Nature. 2011; 477:466–70. [PubMed: 21849977]
19. Sok D, et al. Recombinant HIV envelope trimer selects for quaternary-dependent antibodies targeting the trimer apex. Proc Natl Acad Sci U S A. 2014; 111:17624–9. [PubMed: 25422458]
20. Pancera M, et al. Crystal structure of PG16 and chimeric dissection with somatically related PG9: structure-function analysis of two quaternary-specific antibodies that effectively neutralize HIV-1. Journal of Virology. 2010; 84:8098–110. [PubMed: 20538861]
21. Pejchal R, et al. Structure and function of broadly reactive antibody PG16 reveal an H3 subdomain that mediates potent neutralization of HIV-1. Proc Natl Acad Sci U S A. 2010; 107:11483–8. [PubMed: 20534513]
22. McLellan JS, et al. Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. Nature. 2011; 480:336–43. [PubMed: 22113616]
23. Doria-Rose NA, et al. A Short Segment of the HIV-1 gp120 V1/V2 Region Is a Major Determinant of Resistance to V1/V2 Neutralizing Antibodies. J Virol. 2012
24. Julien JP, et al. Asymmetric recognition of the HIV-1 trimer by broadly neutralizing antibody PG9. Proc Natl Acad Sci U S A. 2013
25. Pancera M, et al. N332-Directed Broadly Neutralizing Antibodies Use Diverse Modes of HIV-1 Recognition: Inferences from Heavy-Light Chain Complementation of Function. PLoS One. 2013; 8:e55701. [PubMed: 23431362]
26. Pancera M, et al. Structural basis for diverse N-glycan recognition by HIV-1-neutralizing V1–V2-directed antibody PG16. Nat Struct Mol Biol. 2013; 20:804–13. [PubMed: 23708607]
27. Sanders RW, et al. A next-generation cleaved, soluble HIV-1 Env trimer, BG505 SOSIP.664 gp140, expresses multiple epitopes for broadly neutralizing but not non-neutralizing antibodies. PLoS Pathog. 2013; 9:e1003618. [PubMed: 24068931]

28. Ross SA, Sarisky CA, Su A, Mayo SL. Designed protein G core variants fold to native-like structures: sequence selection by ORBIT tolerates variation in backbone specification. Protein Sci. 2001; 10:450–4. [PubMed: 11266631]
29. Fazi B, et al. Unusual binding properties of the SH3 domain of the yeast actin-binding protein Abp1: structural and functional analysis. J Biol Chem. 2002; 277:5290–8. [PubMed: 11668184]
30. McLellan JS, et al. Structure of RSV fusion glycoprotein trimer bound to a prefusion-specific neutralizing antibod. Science. 2013 In press.
31. Zhou T, et al. Transplanting supersites of HIV-1 vulnerability. PLoS One. 2014; 9:e99881. [PubMed: 24992528]
32. Do Kwon Y, et al. Crystal structure, conformational fixation and entry-related interactions of mature ligand-free HIV-1 Env. Nat Struct Mol Biol. 2015; 22:522–31. [PubMed: 26098315]
33. Badger J, et al. Structural analysis of a set of proteins resulting from a bacterial genomics project. Proteins. 2005; 60:787–96. [PubMed: 16021622]
34. Tyndall JD, et al. Macrophage migration inhibitory factor covalently complexed with phenethyl isothiocyanate. Acta Crystallogr Sect F Struct Biol Cryst Commun. 2012; 68:999–1002.
35. Pancera M, et al. Structure and immune recognition of trimeric pre-fusion HIV-1 Env. Nature. 2014; 514:455–61. [PubMed: 25296255]
36. Pan R, Gorny MK, Zolla-Pazner S, Kong XP. The V1V2 Region of HIV-1 gp120 Forms a Five-Stranded Beta Barrel. J Virol. 2015; 89:8003–10. [PubMed: 26018158]
37. Doria-Rose NA, et al. Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies. Nature. 2014; 509:55–62. [PubMed: 24590074]
38. Briney BS, Willis JR, Crowe JE Jr. Human peripheral blood antibodies with long HCDR3s are established primarily at original recombination using a limited subset of germline genes. PLoS One. 2012; 7:e36750. [PubMed: 22590602]
39. Briney BS, Willis JR, Hicar MD, Thomas JW 2nd, Crowe JE Jr. Frequency and genetic characterization of V(DD)J recombinants in the human peripheral blood antibody repertoire. Immunology. 2012; 137:56–64. [PubMed: 22612413]
40. Zhu J, et al. Mining the antibodyome for HIV-1-neutralizing antibodies with next-generation sequencing and phylogenetic pairing of heavy/light chains. Proc Natl Acad Sci U S A. 2013; 110:6470–5. [PubMed: 23536288]
41. Shi B, et al. Comparative analysis of human and mouse immunoglobulin variable heavy regions from IMGT/LIGM-DB with IMGT/HighV-QUEST. Theor Biol Med Model. 2014; 11:30. [PubMed: 24992938]
42. Pancera M, et al. Crystal structure of PG16 and chimeric dissection with somatically related PG9: structure-function analysis of two quaternary-specific antibodies that effectively neutralize HIV-1. J Virol. 2010; 84:8098–110. [PubMed: 20538861]
43. Liao HX, et al. Vaccine Induction of Antibodies against a Structurally Heterogeneous Site of Immune Pressure within HIV-1 Envelope Protein Variable Regions 1 and 2. Immunity. 2013; 38:176–86. [PubMed: 23313589]
44. Munro JB, et al. Conformational dynamics of single HIV-1 envelope trimers on the surface of native virions. Science. 2014; 346:759–63. [PubMed: 25298114]
45. Sanders RW, et al. HIV-1 VACCINES. HIV-1 neutralizing antibodies induced by native-like envelope trimers. Science. 2015; 349:aac4223. [PubMed: 26089353]
46. Meffre E, et al. Immunoglobulin heavy chain expression shapes the B cell receptor repertoire in human B cell development. J Clin Invest. 2001; 108:879–86. [PubMed: 11560957]
47. West AP Jr, Diskin R, Nussenzweig MC, Bjorkman PJ. Structural basis for germ-line gene usage of a potent class of antibodies targeting the CD4-binding site of HIV-1 gp120. Proceedings of the National Academy of Sciences of the United States of America. 2012; 109:E2083–90. [PubMed: 22745174]
48. McGuire AT, et al. Engineering HIV envelope protein to activate germline B cell receptors of broadly neutralizing anti-CD4 binding site antibodies. The Journal of experimental medicine. 2013
49. Majeed S, et al. Enhancing protein crystallization through precipitant synergy. Structure. 2003; 11:1061–70. [PubMed: 12962625]

50. Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 1997; 276:307–326.
51. Zhou T, et al. Structural definition of a conserved neutralization epitope on HIV-1 gp120. Nature. 2007; 445:732–7. [PubMed: 17301785]
52. Calarese DA, et al. Antibody domain exchange is an immunological solution to carbohydrate cluster recognition. Science. 2003; 300:2065–71. [PubMed: 12829775]
53. Doria-Rose NA, et al. A new member of the V1V2-directed CAP256-VRC26 lineage that shows increased breadth and exceptional potency. J Virol. 2015
54. Huang J, et al. Broad and potent HIV-1 neutralization by a human antibody that binds the gp41-gp120 interface. Nature. 2014; 515:138–42. [PubMed: 25186731]
55. Scheid JF, et al. Broad diversity of neutralizing antibodies isolated from memory B cells in HIV-infected individuals. Nature. 2009; 458:636–40. [PubMed: 19287373]
56. Chen L, et al. Structural basis of immune evasion at the site of CD4 attachment on HIV-1 gp120. Science. 2009; 326:1123–7. [PubMed: 19965434]
57. Seaman MS, et al. Tiered categorization of a diverse panel of HIV-1 Env pseudoviruses for assessment of neutralizing antibodies. J Virol. 2010; 84:1439–52. [PubMed: 19939925]
58. Kwong PD, et al. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. Nature. 1998; 393:648–59. [PubMed: 9641677]
59. Killikelly A, et al. Thermodynamic signatures of the antigen binding site of mAb 447–52D targeting the third variable region of HIV-1 gp120. Biochemistry. 2013; 52:6249–57. [PubMed: 23944979]
60. Dreyfus C, et al. Highly conserved protective epitopes on influenza B viruses. Science. 2012; 337:1343–8. [PubMed: 22878502]
61. Tharakaraman K, Subramanian V, Cain D, Sasisekharan V, Sasisekharan R. Broadly neutralizing influenza hemagglutinin stem-specific antibody CR8020 targets residues that are prone to escape due to host selection pressure. Cell Host Microbe. 2014; 15:644–51. [PubMed: 24832457]
62. The IMpact-RSV Study Group. Palivizumab a humanized respiratory syncytial virus monoclonal antibody reduces hospitalization from respiratory syncytial virus infection in high-risk infants. Pediatrics. 1998; 102:531–7.
63. Montefiori DC. Measuring HIV neutralization in a luciferase reporter gene assay. Methods Mol Biol. 2009; 485:395–405. [PubMed: 19020839]
64. Shu Y, et al. Efficient protein boosting after plasmid DNA or recombinant adenovirus immunization with HIV-1 vaccine constructs. Vaccine. 2007; 25:1398–408. [PubMed: 17113201]
65. Mastronarde DN. Automated electron microscope tomography using robust prediction of specimen movements. J Struct Biol. 2005; 152:36–51. [PubMed: 16182563]
66. Tang G, et al. EMAN2: an extensible image processing suite for electron microscopy. J Struct Biol. 2007; 157:38–46. [PubMed: 16859925]
67. Guttman M, et al. CD4-Induced Activation in a Soluble HIV-1 Env Trimer. Structure. 2014; 22:974–84. [PubMed: 24931470]
68. Xiang Z, Honig B. Extending the accuracy limits of prediction for side-chain conformations. J Mol Biol. 2001; 311:421–30. [PubMed: 11478870]
69. Petrey D, et al. Using multiple structure alignments, fast model building, and energetic analysis in fold recognition and homology modeling. Proteins. 2003; 53(Suppl 6):430–5. [PubMed: 14579332]
70. Sievers F, Higgins DG. Clustal omega. Curr Protoc Bioinformatics. 2014; 48:3 13 1–3 13 16. [PubMed: 25501942]
71. Xiang Z, Soto CS, Honig B. Evaluating conformational free energies: the colony energy and its application to the problem of loop prediction. Proc Natl Acad Sci U S A. 2002; 99:7432–7. [PubMed: 12032300]
72. Wriggers W, Milligan RA, McCammon JA. Situs: A package for docking crystal structures into low-resolution maps from electron microscopy. J Struct Biol. 1999; 125:185–95. [PubMed: 10222274]

73. Phillips JC, et al. Scalable molecular dynamics with NAMD. J Comput Chem. 2005; 26:1781–802. [PubMed: 16222654]

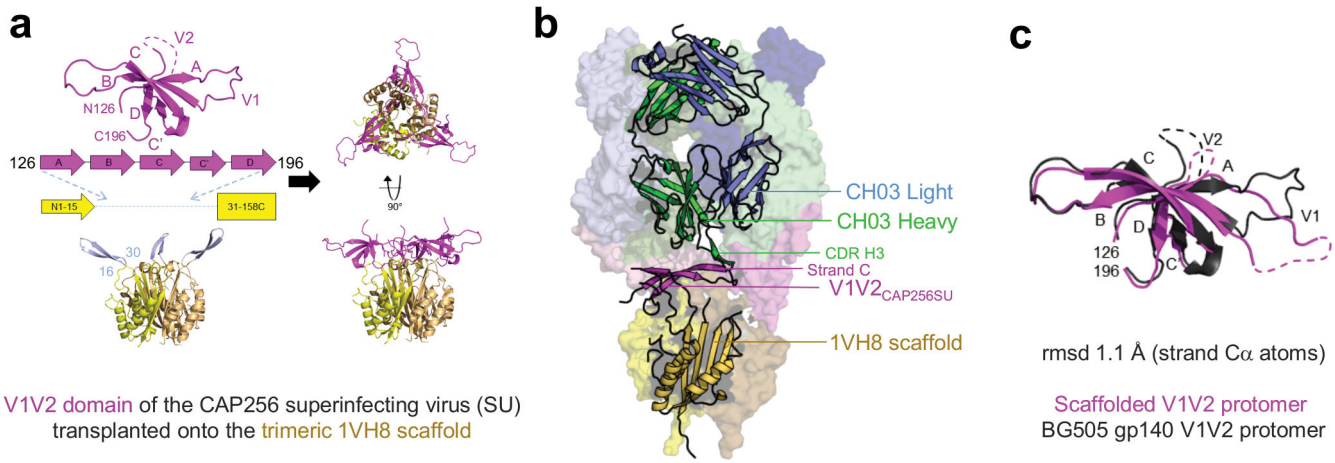

Figure 1.

Design of scaffolded V1V2s and co-crystal structure with bNAb CH03 from donor CH0219. (**a**) Design of trimeric scaffolded V1V2. The trimeric structure of PDB ID 1VH8 used as a scaffold is shown in shades of yellow with the V1V2 domain of BG505 shown in magenta. The residues in 1VH8 replaced with the V1V2 domain are shown in light blue. (**b**) Structure of CH03 (green and blue) bound to a 1VH8-scaffolded (yellow) V1V2 domain (magenta) from the superinfecting strain in CAPRISA donor 256. Three Fabs of CH03 are shown bound to the trimeric 1VH8-$V1V2_{CAP256\text{-}SU}$ scaffold. A single Fab and scaffolded V1V2 protomer are shown in ribbon representation, and the two other trimer-related complexes are shown in surface representations. (**c**) Comparison of 1VH8 scaffolded V1V2 from strain CAP256-SU (magenta) and V1V2 of pre-fusion closed trimer from strain BG505 (black, PDBID 4TVP).

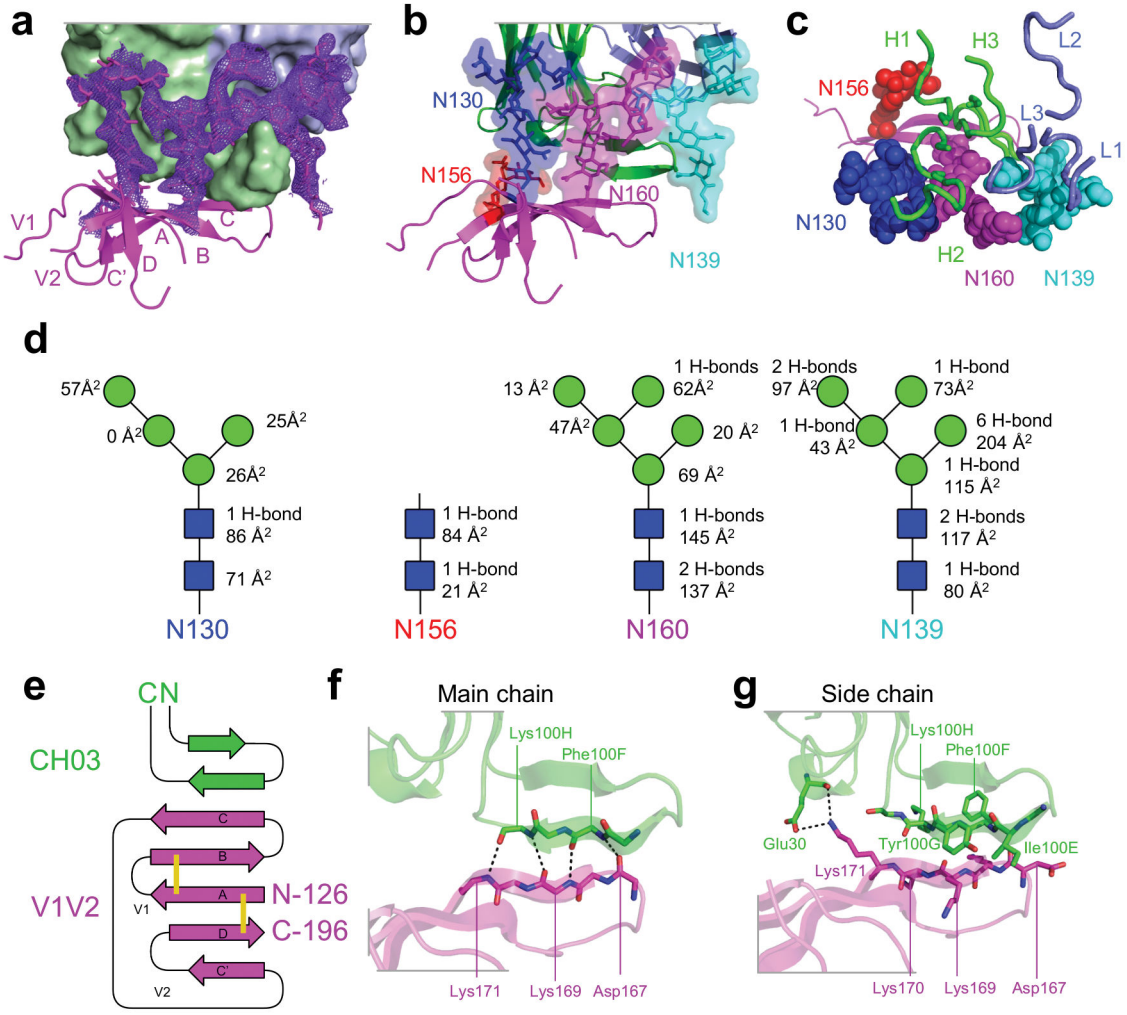

Figure 2.

CH03 V1V2 interactions. (**a**) CH03 is shown as a molecular surface, with V1V2 shown as magenta ribbons. Mannose and *N*-acetylglucosamine residues are shown in stick representation, as are the side chains of Asn 130, 139, 156 and 160. Electron density (2Fo–Fc) is contoured at 1σ and shown as a blue mesh. (**b**) Ribbon representations of CAP256SU V1V2 (magenta), CH03 heavy chain (green) and light chain (light blue). V1V2 glycans are shown in stick and transparent surface representation. (**c**) 90° view of **b**, with CH03 loops shown in ribbon representation and glycans represented as spheres. (**d**) Schematic representation of glycan moieties recognized by CH03. N-acetylglucosamines are shown as blue squares, and mannoses as green circles. Hydrogen bonds to CH03 are listed to the right of the symbols, as is the total surface area buried at the interface between CH03 and each sugar. (**e**) Schematic of the CH03–main-chain interaction with V1V2. Disulphide bonds in V1V2 are shown as yellow sticks. (**f, g**) Ribbon representation of V1V2 (magenta) and CH03 heavy chain (green). Hydrogen bonds are represented by dotted lines. Main-chain interactions are shown in **e**, and side chain interactions in **f**.

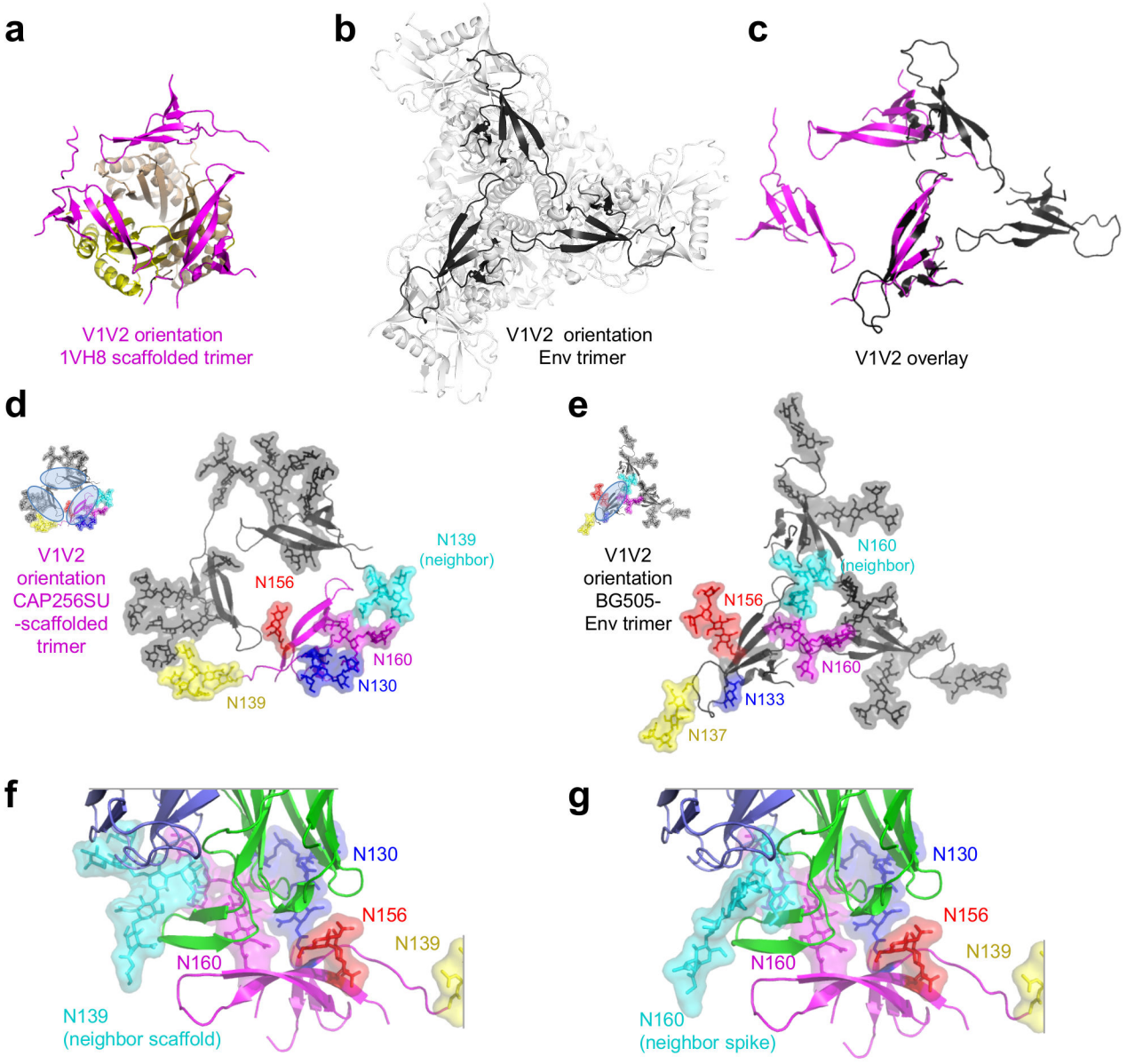

Figure 3.

CH03 epitope in 1VH8-scaffold and trimeric Env contexts. (**a**) V1V2 (magenta) orientation on the trimeric 1VH8 scaffold (yellow). (**b**) V1V2 (black) in trimeric BG505 SOSIP.664 (gray) (**c**) Overlay of V1V2 in scaffold (magenta) and Env (black) contexts, where a single V1V2 has been superimposed. (**d**) 1VH8-scaffolded V1V2 in ribbon representation with glycans as sticks with a transparent surface is shown in color for one protomer, with the other two in gray; the N139 glycan from a neighboring protomer is highlighted in cyan; inset shows binding sites for three CH03 Fabs. (**e**) HIV-1 Env trimeric V1V2 in ribbon representation with glycans in sticks is shown in color for one protomer, with the other two protomers in gray; the N160 glycan from right protomer is highlighted in cyan; inset shows binding site for one CH03 Fab (specific glycan numbering differs based on strain). (**f**) Interaction of CH03 with V1V2 epitope in 1VH8 scaffold context with N139 from the

neighboring protomer highlighted. (**g**) Model of interaction of CH03 with V1V2 epitope in Env trimer context with N160 from the neighboring protomer highlighted. To obtain the model, the CH03 structure was aligned to the BG505.SOSIP trimer through the V1V2 domain and the absent mannose moieties were appended from previous V1V2 structures.

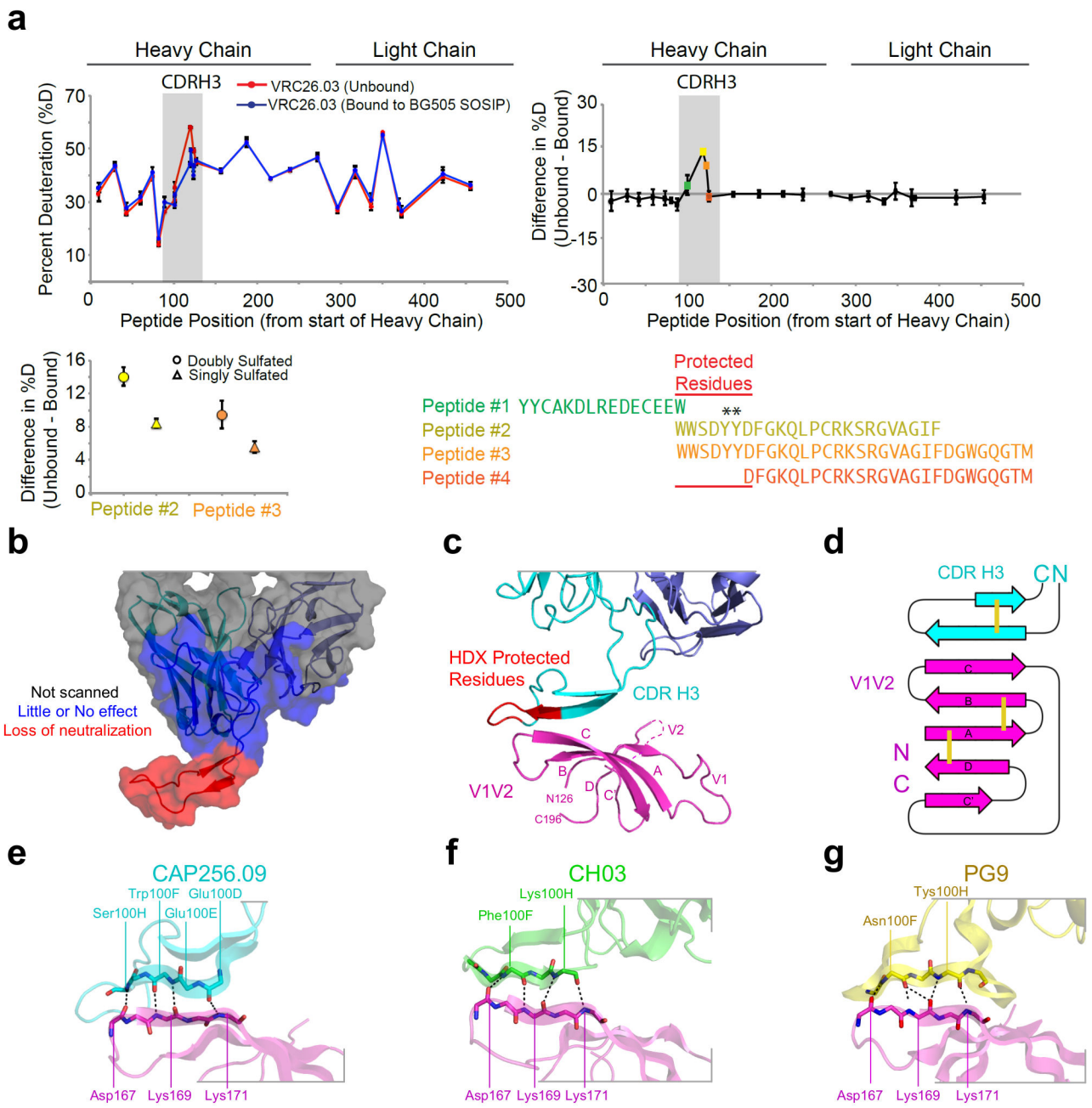

Figure 4.

Modeling of bNAb CAP256-VRC26 bound to Env. (**a**) HDX plots of bound vs unbound CAP256-VRC26 highlight seven CDR H3 residues (red lines) that are protected upon binding to the BG505 SOSIP Env trimer. Specific peptides are colored according to fragment and asterisks indicate sulfated tyrosines within the peptide. (**b**) Structure of CAP256-VRC26 displaying paratope mapping results with regions associated with loss of neutralization shown in red. (**c**) Atomic-level model of CAP256-VRC26 (cyan) bound to V1V2 based on crystal structures of CAP256-VRC26 Fab, HDX, paratope mapping and EM data. CDR H3 residues protected from HDX are shown in red. (**d**) Schematic of the

CAP256-VRC26–main-chain topology with V1V2. Disulfide bonds in V1V2 are shown as yellow sticks. (**e, f, g**) Common mode of recognition by V1V2-directed bNAbs from three donors. A representative antibody is shown for each donor is shown with the atomic-model of CAP256-VRC26.09 (**e**) the crystal structure of CH03–V1V2 (**f**) and the crystal structure of PG9–V1V2 (**g**), with intermolecular H-bond interactions between CDR H3 and V1V2 strand C highlighted with main chain interactions in grey and side chain interactions shown in black.

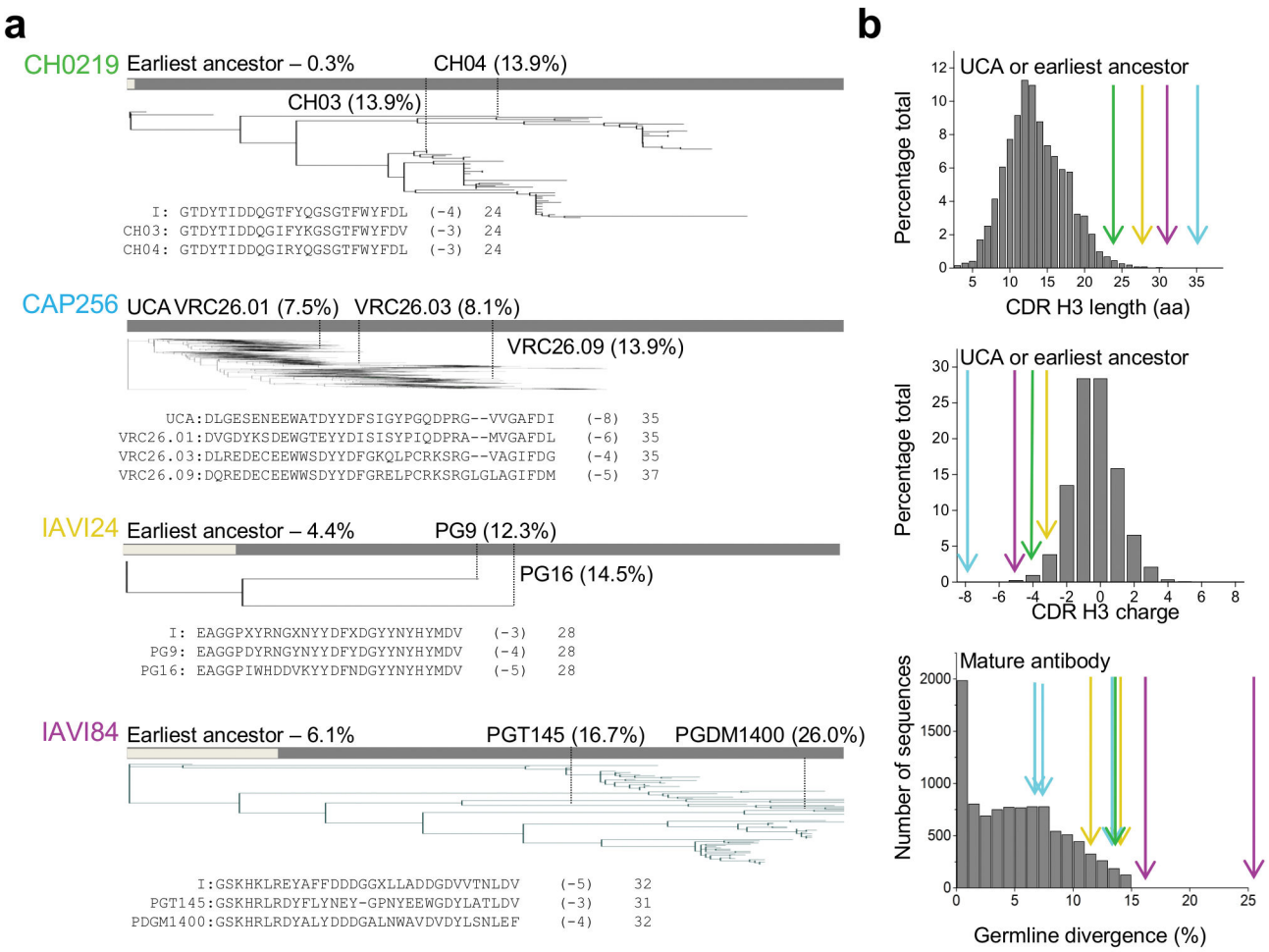

Figure 5.

Properties of UCA or earliest known ancestor of V1V2-directed bNAb lineages. (**a**) For each of the four donors from which V1V2-directed bNAbs have been identified, a phylogenetic tree is displayed with the heavy chain sequences of lineage members. The grey bars denote percentage SHM, with UCA (0% mutated) on the far left, white bars for earliest ancestors, and mature bNAbs labeled. Properties of the CDR H3 are provided for select bNAbs and UCA or earliest known ancestor. The overall charge of the CDR H3 appears in parenthesis followed by the length (Kabat system) (**b**) Histogram of CDR H3 lengths, charge and germline divergence from 9,721 unique and productive human sequences; values for select lineage members are indicated by colored arrows.

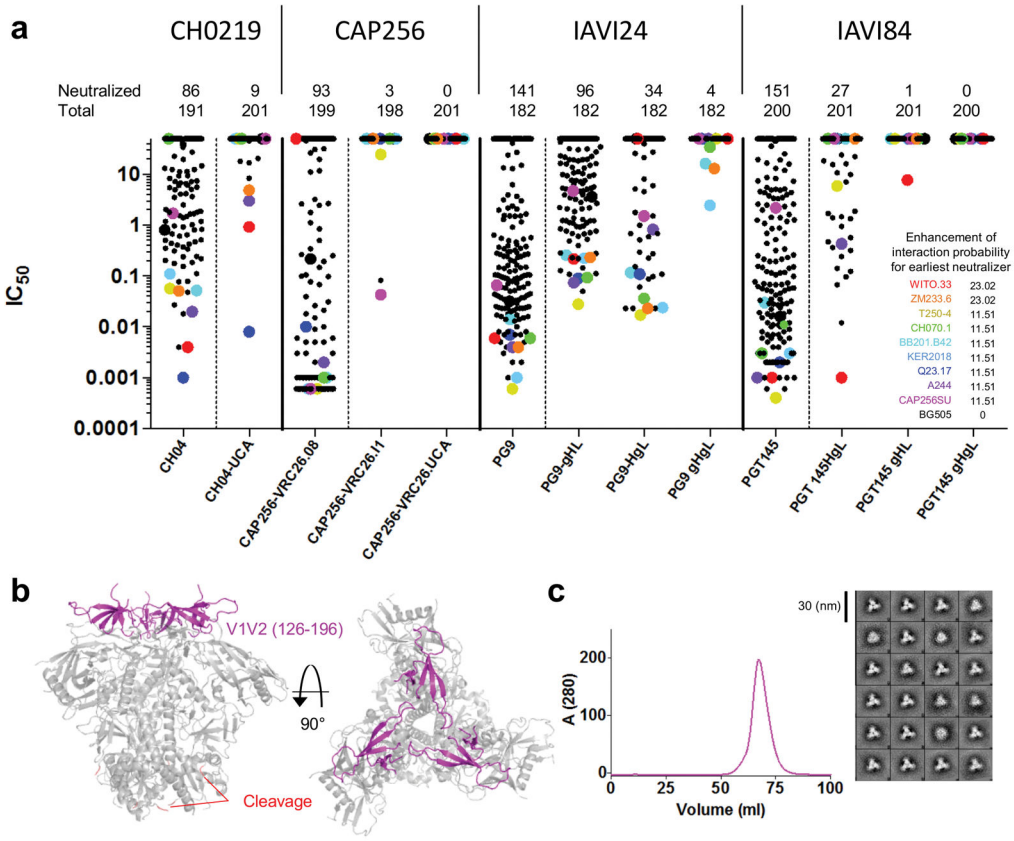

Figure 6.

Inferred ancestor and intermediates of V1V2-directed bNAbs neutralize a common set of HIV-1 isolates. (**a**) Mature and reverted V1V2-directed bNAbs neutralization across ~200 HIV-1 isolates. “Neutralized” represents the number of HIV-1 strains with IC_{50} of less than 50 μg/ml, and “Total” indicates the number of HIV-1 strains tested. Antibodies are organized by donor with mature antibodies to the left of the dashed vertical line and reverted antibodies to the right. IC_{50} for select strains is indicated by enlarged dots colored according to the representative strain shown in the inset. Inset shows rank order of strains neutralized by revertants according to probabilities obtained by frequentist analysis. Enhancement in likelihood of interaction with the earliest neutralizers is provided. Nomenclature of the revertants are as follows: unmutated common ancestor (UCA), reverted V-gene, mature CDR3 (gHgL), early intermediate from next-generation sequencing (I1), CH0219-UCA represents previously inferred common ancestor[15]. See supplementary table 1 for likelihood analysis. (**b**) Design of soluble HIV-1 Env trimers with chimeric V1V2. Residues 126–196 (magenta) of strains found to preferentially interact with germline-reverted V1V2-directed antibodies transferred to the corresponding region of BG505 DS-SOSIP.664, with D368R mutation. (**c**) Gel filtration and negative stain EM (2D class averages) of BG505 SOSIP.664.DS.368R.CAP256SU, a representative chimera.

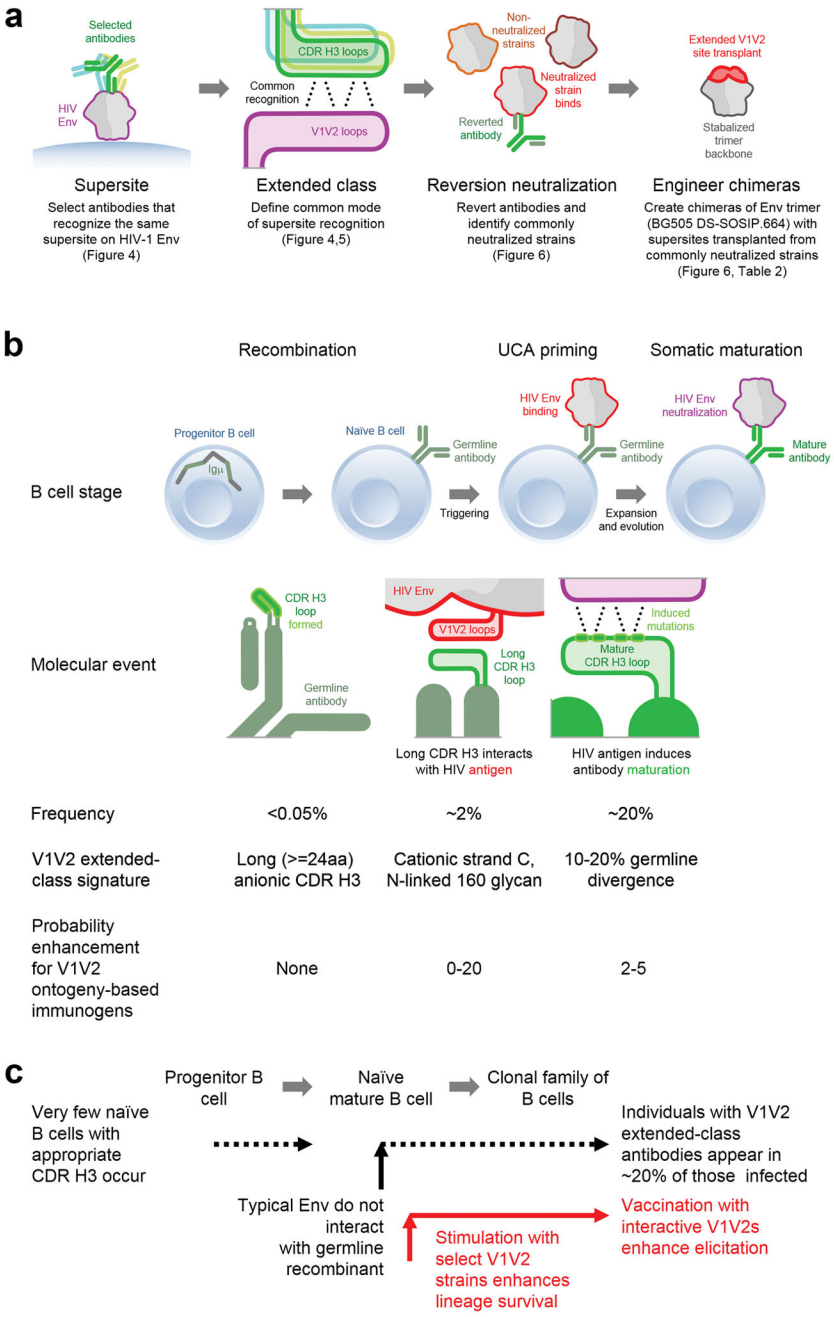

Figure 7.

Ontogeny-based vaccine design. (**a**) A general scheme for the development of ontogeny-based immunogens. (**b**) Different stages of B cell development for V1V2-directed bNAbs and their probabilistic enhancement with ontogeny-based immunogens. The recombination frequency (0.093%) was estimated based on the percentage of long (equal to or greater than 24aa) and anionic (net charge of -3 or lower) CDRH3 present in 9,721 unique and productive human sequences. Probability enhancements are shown for a single immunogen; substantial additional enhancement could be gained through the use of immunogen cocktails

(see Supplementary Table 1). (**c**) Schematic of the B cell ontogeny for the extended class of V1V2-directed bNAbs. Strategies for vaccine enhancement are highlighted in red.

Table 1

Data collection and refinement statistics (molecular replacement)

	CH03-1VH8-V1V2$_{SU}$	CH04-1VH8-V1V2$_{A244}$
Data collection		
Space group	C2	$P6_3$
Cell dimensions		
a, *b*, *c* (Å)	162.69, 98.21, 170.63	116.72, 116.72, 249.56
α, β, γ (°)	90, 112.82, 90	90, 90, 120
Resolution (Å)	50.00-3.10 (3.17-3.10)	50.00-4.20 (4.32-4.20)
R_{merge}	18.4 (70.9)	9.7 (76.9)
$I/\sigma I$	7.3 (1.7)	19.2 (1.8)
Completeness (%)	99.7 (98.4)	95.8 (77.6)
Redundancy	3.4 (2.2)	9.6 (5.7)
Refinement		
Resolution (Å)	41.0-3.10	50.0-4.2
No. reflections	44633 (4164)	13064 (777)
R_{work}/R_{free}	21.3/25.8	24.2/28.5
No. atoms (total)	15133	9388
Protein	14435	9066
Carbohydrate	779	322
B-factors (Å)^2		
Protein	67	81
Carbohydrate	94.5	96
R.m.s. deviations		
Bond lengths (Å)	0.004	0.006
Bond angles (°)	0.8	1.34

Values in parentheses are for highest-resolution shell.

One crystal was used to measure the data for each complex.

Table 2

and Binding of Selected V1V2 Strains

		CH0219				CAP256					IAVI24					IAVI84				
lade	**Assay**	**CH01**	**CH04**	***CH-UCA1***	***CH-UCA2***	**CAP256.08**	**CAP256.25**	***CAP256-I1***	***CAP256.09 gHgL***	***CAP256.UCA***	**PG9**	**PG16**	***PG9 gHL***	***PG9 HgL***	***PG9 gHgL***	**PGT145**	**PGT142**	***PGT145 gHL***	***PGT145 HgL***	***PGT145 gHgL***
B	ELISA EC_{50}	+++	+++	+++	+++	-	-	-	-	-	+++	+++	++	+++	+	+++	+++	++	+++	-
	Neut IC_{50}	0.013	0.004	*3.78*	*0.923*	>50	>50	*>50*	*>50*	*>50*	<.023	<.023	*0.217*	*>50*	*>50*	0.0003	0.0003	*7.76*	*0.001*	*>50*
C	ELISA EC_{50}	+++	+++	+	-	+++	+++	++	++	-	+++	+++	+++	+++	+++	+++	+++	-	++	-
	Neut IC_{50}	0.074	0.051	**38*	**4.9*	0.001	*0.0003*	*>50*	*1.98*	*>50*	<.023	<.023	*0.231*	*<.023*	*13.1*	0.011	0.018	*>50*	*>50*	*>50*
AG	ELISA EC_{50}	++	+++	-	-	+++	+++	+++	+++	-	+++	+++	+++	+++	++	+++	+++	-	+	-
	Neut IC_{50}	0.042	0.057	*>50*	*>50*	0.0003	0.0003	*24.8*	*0.0003*	*>50*	0.004	0.0005	*0.028*	*0.017*	*>50*	0.0002	0.003	*>50*	*5.93*	*>50*
BC	ELISA EC_{50}	+	+	-	-	-	+	-	-	-	++	++	+	+	-	+	+	-	-	-
	Neut IC_{50}	>50	>50	*>50*	*>50*	0.039	*0.0003*	*>50*	*>50*	*>50*	0.006	0.002	*0.093*	*0.036*	*34.6*	0.001	0.015	*>50*	*>50*	*>50*
A	ELISA EC_{50}	+++	+++	++	++	+++	+++	-	++	-	+++	+++	+++	+++	++	+++	+++	-	+	-
	Neut IC_{50}	0.126	0.052	*>50*	*>50*	0.001	*0.0003*	*>50*	*>50*	*>50*	0.014	0.003	*0.256*	*0.117*	*16.4*	0.03	0.045	*>50*	*>50*	*>50*
A	ELISA EC_{50}	+++	+++	+	+	+++	+++	+	+	-	+++	+++	++	++	+	+++	+++	-	+	-
	Neut IC_{50}	0.159	0.11	*>50*	*>50*	0.0003	*0.0003*	*>50*	*0.018*	*>50*	0.001	<.0006	*0.223*	*0.024*	*2.45*	0.003	0.0003	*>50*	*>50*	*>50*
A	ELISA EC_{50}	+++	+++	+++	+++	-	+	-	-	-	+++	+++	+++	+++	++	+++	+++	-	-	-
	Neut IC_{50}	0.008	0.001	*0.013*	*0.008*	0.01	*0.0003*	*>50*	*>50*	*>50*	0.007	0.002	*0.09*	*0.1*	*>50*	0.002	0.0008	*>50*	*>50*	*>50*
AE	ELISA EC_{50}	+++	+++	++	++	++	++	-	-	-	++	++	+	+	+	+++	+++	-	+	-
	Neut IC_{50}	<.02	0.02	**7.9*	**3*	0.002	0.092	*>50*	*>50*	*>50*	0.004	0.001	0.074	0.831	*>50*	0.0003	0.0003	*>50*	*>50*	*>50*
C	ELISA EC_{50}	++	++	-	-	+++	+++	+++	+++	+	+++	+++	++	++	++	+++	+++	-	-	-
	Neut IC_{50}	2.13	1.72	*>50*	*>50*	0.003	0.0005	*0.04*	*0.01*	***>50*	0.04	0.005	*>50*	*>50*	*>50*	1.25	0.019	*>50*	*>50*	*>50*

ned from previously published results[15].

** The CAP256SU strain is partially neutralized by the CAP256.UCA but not well enough to obtain an IC_{50}