# Long antibody HCDR3s from HIV-naïve donors presented on a PG9 neutralizing antibody background mediate HIV neutralization

Jordan R. Willis[a,b,c,1], Jessica A. Finn[b,c,d], Bryan Briney[c,d,2], Gopal Sapparapu[c,e], Vidisha Singh[c], Hannah King[c], Celia C. LaBranche[f], David C. Montefiori[f], Jens Meiler[a,b,g,h,3], and James E. Crowe Jr.[a,c,d,e,3]

[a]Chemical and Physical Biology Program, Vanderbilt University, Nashville, TN 37232; [b]Center for Structural Biology, Vanderbilt University, Nashville, TN 37232; [c]Vanderbilt Vaccine Center, Vanderbilt University, Nashville, TN 37232; [d]Department of Pathology, Microbiology, and Immunology, Vanderbilt University, Nashville, TN 37232; [e]Department of Pediatrics, Vanderbilt University, Nashville, TN 37232; [f]Department of Surgery, Duke University Medical Center, Durham, NC 27710; [g]Department of Chemistry, Vanderbilt University, Nashville, TN 37232; and [h]Institute for Chemical Biology, Vanderbilt University, Nashville, TN 37232

Edited by David Baker, University of Washington, Seattle, WA, and approved February 26, 2016 (received for review September 16, 2015)

**Development of broadly neutralizing antibodies (bnAbs) against HIV-1 usually requires prolonged infection and induction of Abs with unusual features, such as long heavy-chain complementarity-determining region 3 (HCDR3) loops. Here we sought to determine whether the repertoires of HIV-1–naïve individuals contain Abs with long HCDR3 loops that could mediate HIV-1 neutralization. We interrogated at massive scale the structural properties of long Ab HCDR3 loops in HIV-1–naïve donors, searching for structured HCDR3s similar to those of the HIV-1 bnAb PG9. We determined the nucleotide sequences encoding $2.3 \times 10^7$ unique HCDR3 amino acid regions from 70 different HIV-1–naïve donors. Of the 26,917 HCDR3 loops with 30-amino acid length identified, we tested 30 for further study that were predicted to have PG9-like structure when chimerized onto PG9. Three of these 30 PG9 chimeras bound to the HIV-1 gp120 monomer, and two were neutralizing. In addition, we found 14 naturally occurring HCDR3 sequences that acquired the ability to bind to the HIV-1 gp120 monomer when adding 2- to 7-amino acid mutations via computational design. Of those 14 designed Abs, 8 neutralized HIV-1, with $IC_{50}$ values ranging from 0.7 to 98 μg/mL. These data suggest that the repertoire of HIV-1–naïve individuals contains rare B cells that encode HCDR3 loops that bind or neutralize HIV-1 when presented on a PG9 background with relatively few or no additional mutations. Long HCDR3 sequences are present in the HIV-naïve B-cell repertoire, suggesting that this class of bnAbs is a favorable target for rationally designed preventative vaccine efforts.**

HIV | neutralizing antibodies | molecular conformation | protein design

Elicitation of broadly neutralizing Abs (bnAbs) against HIV type 1 (HIV-1) is one of the greatest challenges in modern vaccinology (1). A bnAb response occurs at various levels in most HIV-1–infected individuals and is extraordinarily broad and potent in a minor subset of subjects (2, 3). These bnAbs typically arise only a year or more after infection and peak at ∼3–4 y after infection (4). Recent advances in donor selection from cohorts of chronically infected patients, novel screening methods, and Ab isolation technologies have allowed identification of dozens of new bnAbs specific to the envelope glycoproteins gp120/gp41 of HIV-1 (5). Although bnAbs of interest have been isolated using different strategies, the genes encoding these bnAbs often share certain distinctive and unusual features such as excessively large amounts of somatic mutation or very long heavy-chain complementarity-determining region 3 (HCDR3) structures (6). The requirement for these unusual Ab genetic features as the basis of formation of HIV-1 bnAbs is an obstacle to vaccination, because conventional vaccines fail to induce a high frequency of Abs with such characteristics.

Although studies based on bnAb isolation have revealed new targets for immunogen design and have been useful for experimental therapeutic efficacy (7), the design of an effective vaccine against HIV-1 has remained elusive (8). There is a structural class of bnAbs represented by PG9 and PG16 whose members have long protruding anionic HCDR3 structures that form beta-sheet hammerheads and target the V1/V2 epitope (9–11). These bnAbs can be very broad and potent, neutralizing up to ∼80% of viral variants tested (12). This class of Abs often has lower levels of somatic mutation relative to other bnAbs, making them attractive targets for a rational vaccine design strategy. However, the genetic basis for formation of these types of long HCDR3s is poorly understood. PG9 was isolated from an HIV-1–infected individual, but here we sought to determine whether any Ab sequences in the HIV-1–naïve donor B-cell repertoire encode long HCDR3 structures that can bind and neutralize HIV-1 when presented in the context of the naturally occurring PG9 antibody. The presence of such Abs in the HIV-1–naïve repertoire might suggest optimal paratope targets for vaccine B-cell priming by rationally designed HIV-1 envelope (Env) antigens. To search for HIV-1–reactive HCDR3 sequences in the HIV-1–naïve repertoire, we developed an experimental and bioinformatics pipeline that coupled next-generation sequencing with 3D structure prediction and functional design and then

## Significance

When HIV neutralizing antibody (Ab) responses occur, they are often mediated by Abs with exceptional levels of somatic mutation. An exception are HIV neutralizing Abs that feature long heavy-chain complementarity-determining region 3 (HCDR3) regions, such as the Ab PG9. Relative to many other HIV broadly neutralizing Abs (bnAbs), PG9 has fewer somatic mutations than most potent bnAbs. Here we used deep-sequencing and computational methods to identify a panel of HCDR3 sequences in HIV-naïve donors that mediated binding and neutralization of HIV by mimicking the bnAb PG9 long HCDR3 region when expressed in the context of the rest of the PG9 antibody sequence. Thus, it may be possible for structure-based vaccine design efforts to engage long HCDR3s available in the HIV-naïve Ab repertoire and elicit a neutralizing response.

Author contributions: J.R.W., J.A.F., B.B., J.M., and J.E.C. designed research; J.R.W., J.A.F., B.B., G.S., V.S., H.K., C.C.L., and D.C.M. performed research; J.A.F. contributed new reagents/analytic tools; J.R.W., J.A.F., C.C.L., D.C.M., and J.M. analyzed data; and J.R.W., J.M., and J.E.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]Present address: Department of Immunology and Microbial Science, The Scripps Research Institute, La Jolla, CA 92037.

[2]Present address: Department of Molecular Biology, The Scripps Research Institute, La Jolla, CA 92037.

[3]To whom correspondence may be addressed. Email: jens.meiler@vanderbilt.edu or james.crowe@vanderbilt.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1518405113/-/DCSupplemental.

validated the findings with experimental characterization (Fig. 1). We found two naturally occurring HCDR3 sequences from 2 different donors of 70 studied that were predicted to adopt a PG9-like hammerhead conformation and were able to bind and neutralize PG9-susceptible viruses. In addition, we used computational design to mimic the process of maturation by somatic mutation of HCDR3 sequences from the HIV-1–naïve repertoire that were predicted to adopt a PG9-like hammerhead conformation. We identified two to seven mutations in eight different HCDR3 sequences that facilitated neutralization of HIV when grafted on a PG9 Ab background.

## Results

To obtain a large number of long HCDR3 sequences for this analysis, we collected white blood cells from the leukoreduction filters of 70 different HIV-1–negative blood donors from the Nashville, TN, American Red Cross. Using a primer design that amplified the HCDR3 sequence from the Ab heavy chain (*SI Appendix*, Fig. S1), we generated amplicons for sequence analysis on the Illumina HiSeq platform by RT-PCR. This approach generated $2.3 \times 10^7$ unique HCDR3 amino acid sequences (*SI Appendix*, Fig. S2). The mean length for the HCDR3 sequences was $16.4 \pm 4.0$ amino acids. Because the amplicon sequence started from the framework 3 region, only assignments of $V_H$ families and not $V_H$ genes could be resolved unambiguously. However, the distribution of gene assignments (*SI Appendix*, Fig. S3) matched previous observations of gene distribution in the repertoire of healthy donors using analysis of full-length Ab variable gene sequences (13). *IGHD3-3* and *IGHJ6* gene segments were used more frequently as the length of the HCDR3 increased.

**Structural Modeling Provides a Rapid Heuristic to Predict PG9 Mimicry.** To facilitate large-scale structural modeling of HCDR3s, we only considered sequences with HCDR3 loops that were the same length as PG9 (30 amino acids). This approach identified 26,917 unique sequences to be examined for PG9-like structure and activity. For efficient identification of the sequences most likely to exhibit a PG9-like HCDR3 structure, we developed a testing heuristic we designated a position-specific structure scoring matrix (P3SM) that used Rosetta energy scores for a relatively small number of sequences and extrapolated those predictions to the whole sequence pool. In this way, we could explore a large sequence library to identify a subset of sequences with a higher likelihood of mimicking PG9 structure, even though it was not feasible computationally to predict the structure of all members using Rosetta. First, we obtained 4,000 randomly selected sequences from the pool of 26,917 HCDR3s with 30-amino acid length. Next, the naturally occurring HCDR3 sequences were threaded over the wild-type PG9 (PG9*wt*) structure after removing the Env V1/V2 antigen from the known Ab–antigen structure [Protein Data Bank (PDB) ID code 3U4E (11); Movie S1]. We then energetically minimized the sequence to check the "tolerance" for maintaining the PG9*wt* hammerhead structure (Movie S2). HCDR3 sequences were evaluated for structural mimicry of PG9*wt* by how well they retained the PG9*wt* topology, measured as rmsd to PG9*wt*, and the sequence tolerance to the PG9 topology, measured as the Rosetta energy (Fig. 2*A*). We found three different distributions for the energetically minimized sequences: (*i*) sequences that retained PG9*wt* HCDR3 topology but had unfavorable energies (Fig. 2*B*); (*ii*) sequences that had favorable energies but did not retain PG9*wt* topology (Fig. 2*C*); or (*iii*) sequences that retained both HCDR3 topology and favorable energy scores (Fig. 2*D*).

Rosetta allows examination of scoring terms on a per-residue basis (14). Therefore, we filled the P3SM with an average energy for each amino acid identity seen in the naturally occurring sequences. Fig. 2*E* shows the P3SM analysis results as a heat map, where each amino acid identity was assigned an average energy as calculated by Rosetta. For example, positions $100_C$ and $100_D$ preferred glycine, as these positions have a very narrow range of



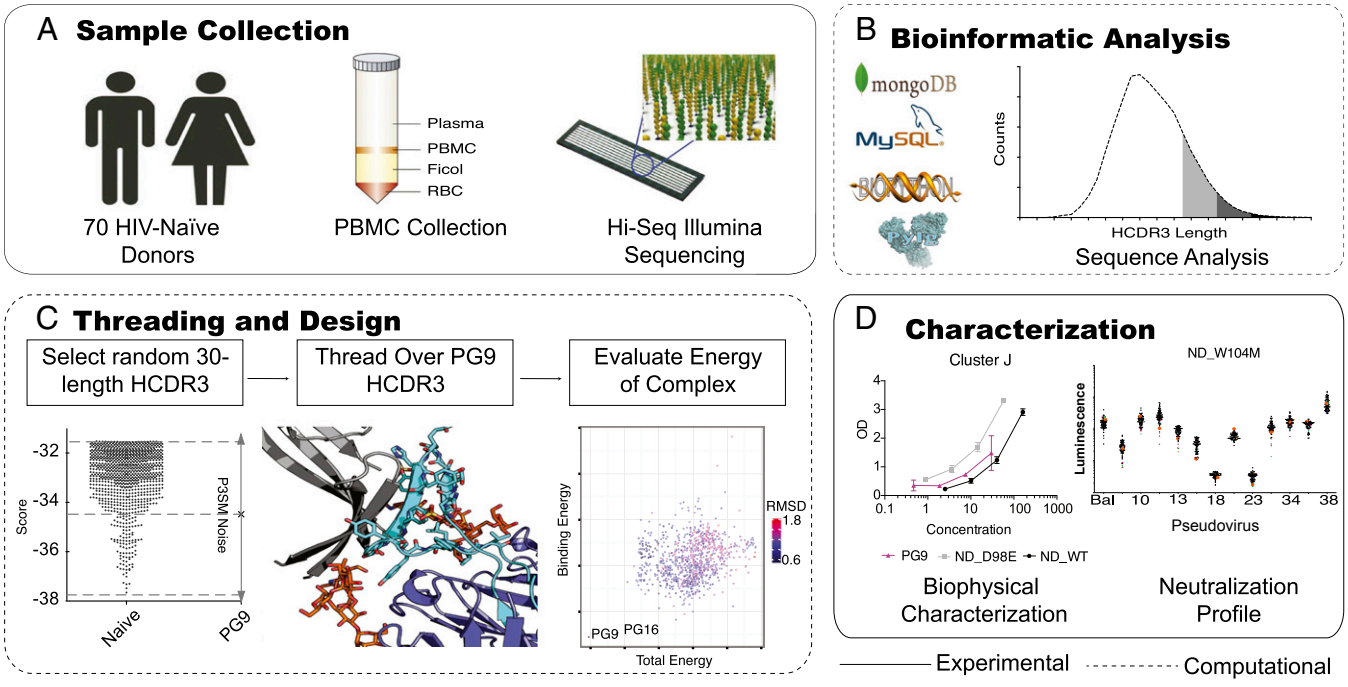


**Fig. 1.** Summary of methodology. The methodology used was a combination of computational (dashed borders) and experimental procedures (solid borders). (*A*) HIV-naïve donor blood was collected from 70 adult donors and the HCDR3 genetic repertoire was sequenced on the Illumina HiSeq platform. PBMC, peripheral blood mononuclear cells. (*B*) The raw sequences were reconstructed and analyzed against germ-line databases using custom software. The sequences were parsed and stored in optimized databases to handle the large quantity of Ab sequences. (*C*) HCDR3 sequences were chosen by length and tested for mimicry of the PG9 HCDR3 structure using the Rosetta software suite. Iterative rounds of minimization, docking, and design, followed by rigorous statistical analysis, allowed for a robust prediction of potential candidates from the HIV-naïve donor repertoire that may bind or neutralize HIV. (*D*) A tractable number of sequences was synthesized and tested experimentally through biophysical characterization and binding and neutralization studies against HIV.

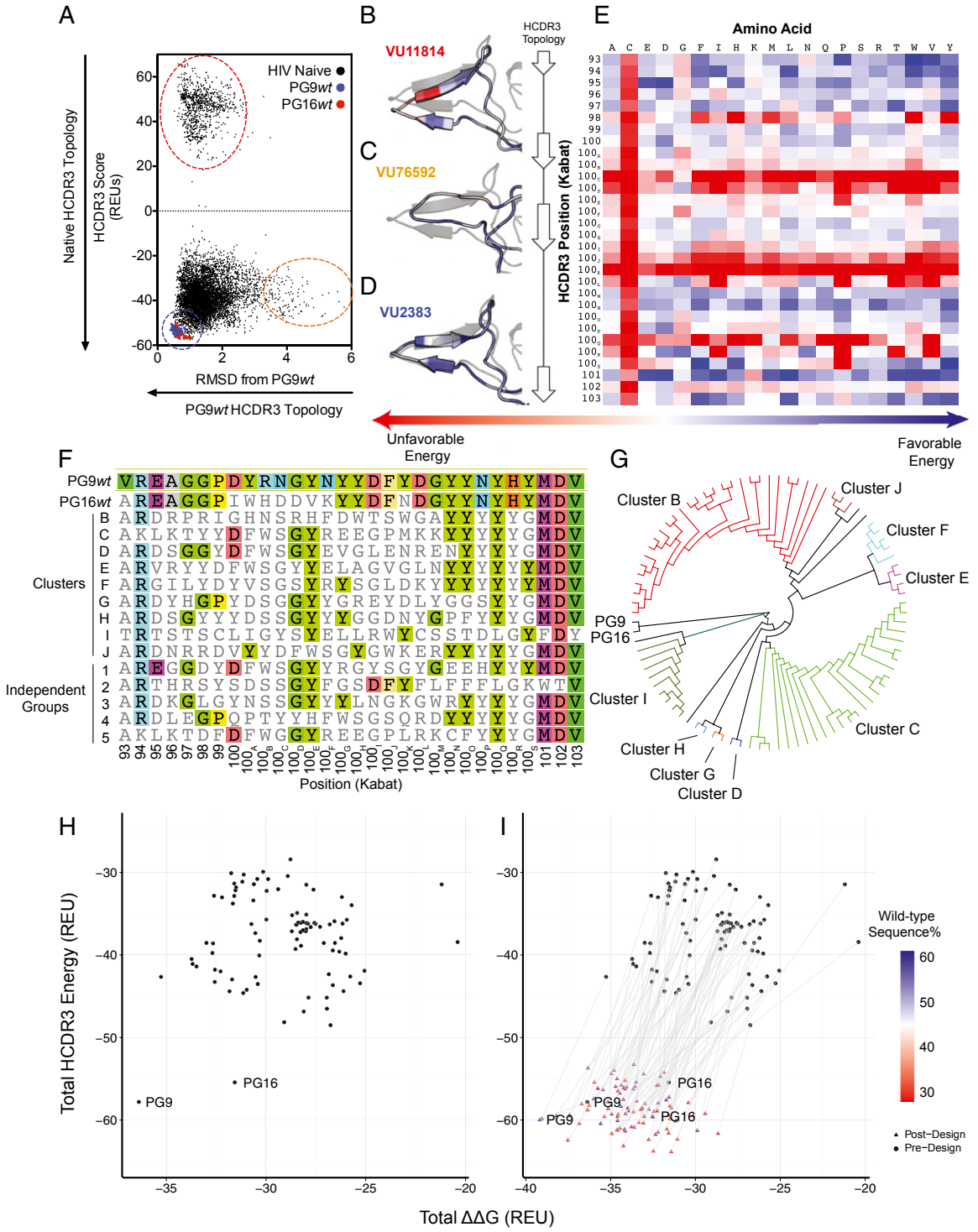


**Fig. 2.** Screening and redesign of HIV-1–naïve donor HCDR3 sequences using Rosetta. (*A*) Four thousand randomly selected sequences that encode HCDR3s that are 30 amino acids in length were selected from next-generation amplicon sequence libraries generated from the blood samples of 70 HIV-naïve donors. The predicted amino acid sequences were threaded onto the PG9*wt* backbone structure (PDB ID code 3U4E) and their tolerance to PG9*wt* as a measure of overall energy (*y* axis) and retention of PG9*wt* HCDR3 structure (*x* axis) were predicted. Abs from HIV-1–naïve donors are shown in black. PG9*wt* and PG16*wt* are shown in blue and red, respectively. (*B–D*) Three outcomes were produced, as follows. (*B*) Sequences that retained the HCDR3 topology but increased in energy. (*C*) Sequences that lowered in energy but could not retain the HCDR3 topology. (*D*) Sequences that could retain the topology and had a favorable energy score. The color of each amino acid corresponds to a position on a blue–red scale for favorable to unfavorable energy, as assessed by Rosetta. (*E*) Each of the sequences from *A* constructs a position-specific structure-scoring matrix (P3SM), where each of the HCDR3 positions is shown on the *y* axis and each amino acid identity is shown on the *x* axis. The average energy of each amino acid for each position fills the matrix. The topology of the HCDR3 loop is shown as a reference. (*F*) The top 100 sequences clustered into nine groups, designated clusters B–J, and five additional sequences that did not fit a group, "independent groups" 1–5 (designated IG1–5). A consensus sequence for each group is aligned with PG9. (*G*) The dendrogram for each of the sequences shows the clustering. (*H*) Scoring of the top 100 sequences assessed by Rosetta is shown as a function of binding energy (*x* axis) and total energy (*y* axis). (*I*) The energetic gap was recovered by minimal redesign of *wt* sequences (circles) and produced sequences with a few mutations (triangles). The *wt* sequences were characterized for sequence recovery (red–blue scale).

torsional angles that accommodate the hinge region of the hammerhead.

Next, all 26,917 sequences with 30-amino acid length were rank-ordered by their P3SM score. We ranked the P3SM scores assuming that the PG9*wt* sequence should be the top-scoring sequence tested. We found that PG9*wt* ranked 92nd (0.4%), scoring 3.82 ROSETTA energy units (REUs) worse than the best sequence (*SI Appendix*, Fig. S5*C*). We rationalized that PG9*wt* should have a lower energy than all sequences from the naïve repertoire and therefore introduced 3.82 REUs as a minimum error margin of the P3SM score. To avoid excluding potential hit sequences, we selected the top 1,000 HCDR3 sequences that scored within ±3.82 REUs of PG9*wt* for further analysis. These sequences were submitted to a more accurate ROSETTA energy evaluation protocol that was too time-consuming to apply to all 26,917 sequences (*SI Appendix*, Fig. S6) (15, 16). In this protocol, the full complex of PG9*wt* bound to the HIV-1 CAP45 strain V1/V2 scaffold (PDB ID code 3U4E) was used in modeling. Glycans at positions 156 and 160 (lab-adapted HIV strain HXBc2 numbering) were reconstituted for evaluation by the ROSETTA scoring function (*SI Appendix*, Fig. S4). To include parts of the scoring function that selected for other HCDR3 sequences in which PG9 and PG16 were in the top 10% of the scores but not necessarily the top-scoring, we found that scores affiliated with glycan modification, HCDR3 contribution to binding energy, and the shift in Cα rmsd also ranked PG9 and PG16 HCDR3s favorable. Because these numbers are not directly comparable, the Z-score metric was introduced to rescale each score and combine them into one composite metric. Using this weighted Z-score metric, 100 models were generated for each sequence (*SI Appendix*, *Materials and Methods* and Fig. S7). The top 100 sequences from the weighted Z-score metric were used for further analysis. Using Clustal W2 (17), we performed a multiple sequence analysis and subsequent phylogeny construction to see whether sequences identified in the rank order were related. Indeed, the sequences clustered to nine unique groups (clusters B–J, containing ≥2 members) and five independent group clusters with a single member (designated IG1–5, Fig. 2 *F* and *G* and *SI Appendix*, Table S1).

**Antibody Design Compensates for Incomplete PG9 Mimicry by Naturally Occurring HCDR3 Sequences.** As seen in Fig. 2*H*, an energetic gap occurred between each of the HIV-1–naïve donor sequences and PG9/16 when comparing thermodynamic stability in the PG9*wt* conformation and predicted binding affinity to the HIV-1 CAP45 strain V1/V2 scaffold. This gap generally was caused by the effect of a few amino acids in the naturally occurring *wt* sequence that scored poorly when adopting the PG9*wt* topology. This gap is expected, as no HCDR3 sequence from the HIV-naïve repertoire was optimized to adopt PG9*wt* topology via somatic mutations. To mimic this maturation process and remove the energetic gap, we redesigned the starting sequences of the HCDR3s to simultaneously optimize stability and binding. To limit the number of mutations introduced by ROSETTADESIGN, an energetic bonus to the starting sequence was introduced (18). As a result, each of the HCDR3 variants mutated between 30% and 60% of the amino acids (Fig. 2*I*). We visually inspected mutations proposed by ROSETTA and reverted some that appeared not motivated by structural context. This process is needed, as the ROSETTA scoring function is limited in accuracy by its pairwise decomposability and a distance cutoff when considering interactions. Further, the design process is stochastic and might terminate in local energetic minima. The visual inspection and triage ensure that only necessary mutations are introduced (*SI Appendix*, Fig. S8).

**HIV-Naïve Donor HCDR3 Sequences on a PG9 Background Enable HIV gp120 Binding.** The top-scoring *wt* sequence from each cluster and a combination of designed sequences provided a total of 84 HCDR3 predicted PG9-like sequences that we synthesized and cloned into a plasmid encoding PG9*wt*, thus replacing the PG9*wt* HCDR3 with an HIV-1–naïve donor HCDR3 sequence (*SI Appendix*, Table S2). These expression experiments resulted in predicted PG9-like chimeras in which the HCDR3 from the naïve or designed sequences was placed onto a PG9 background. The PG9 background retains 17 mutations in the $V_H$ region and 13 mutations in the $V_L$ region. We screened each of the variant Ab supernatants for level of expression of IgG and for binding to HIV-1 using a mixture of eight gp120 monomers as antigen in an ELISA binding assay. Ab variants were named by clone number and how many mutations away from the *wt* sequence each contained. For example, VU2383_5MUTD is Vanderbilt University (VU) site HIV-1–naïve donor *wt* clone 2383 derived from an HCDR3 sequence with five mutations; the letter D indicates that this was the fourth 5MUT variant in sequence that was tested in this group of antibody variants. If the variant Ab clone expressed IgG and that Ab bound to Env antigens in the gp120 ELISA at the maximum endpoint titer, it was considered further (*SI Appendix*, Fig. S9 and Table S3). Fourteen of the 84 Abs failed to express recombinantly, and 32 of the 84 Abs expressed but did not bind at the maximum supernatant titer. These findings resulted in the identification of 30 predicted PG9-like HCDR3 variant sequences that were evaluated for binding and neutralization on a larger scale (Fig. 3*A*). Of the 30 sequences evaluated, 16 bound at least one of the gp120 monomers at a concentration less than 100 μg/mL. We chose 100 μg/mL as our cutoff for binding, as this level was found previously to be less than the biologically relevant cutoff of 1 μM $K_D$ necessary to produce an immune response (19–21). Of those, 2 were based on completely *wt* HIV-1–naïve donor HCDR3 sequences (VU4322, VU65128) and 14 clones had two to seven altered amino acids that were designed by ROSETTA. As expected, the HIV-1 gp120 monomers recognized most frequently by the HCDR3 variant Abs were derived from HIV-1 strains BaL.01 and ZM109F.PB, which were determined previously to bind PG9*wt* strongly (11). The Ab with the tightest binding (VU28693_5MUT, with an $EC_{50}$ of 1.03 μg/mL) had five mutations (17% mutated) compared with the originally isolated HIV-1–naïve donor HCDR3 sequence. The Ab based on a *wt* HIV-1–naïve donor HCDR3 sequence with the tightest binding that had no designed mutations, designated VU42232, bound to HIV-1 strain ZM109F.PB gp120 with an $EC_{50}$ of 3.75 μg/mL. The Ab that exhibited the broadest reactivity for diverse strains was VU28693_2MUTB, which bound four out of the eight gp120 variants tested (Fig. 3*A* and *SI Appendix*, Table S4).

**HIV-Naïve Donor HCDR3 Sequences on a PG9 Background Enable HIV Neutralization.** For neutralization screening, Abs VU65128 and VU4946, with *wt* HCDR3 sequences, neutralized at an $IC_{50}$ of 17.7 and 46.1 μg/mL, respectively. Interestingly, although VU4946 neutralized, it did not bind any of the gp120 monomers tested, a phenomenon that has been observed for this class of Ab (11, 22). Conversely, VU42232 bound strain BaL.01 and ZM109F.PB gp120 at 3.4 and 3.2 μg/mL, respectively, but did not neutralize any of the variants tested. A group of nine minimally designed HCDR3 sequences neutralized at least one of the variants tested at concentrations ranging from 0.7 to 98 μg/mL. The Ab with the most potent neutralizing activity for any clone tested was VU2383_3MUTD, which neutralized the HIV-1 strain CAP45.2.00.G3 with an $IC_{50}$ of 0.7 μg/mL. The Ab with the broadest neutralizing activity for diverse strains was VU30400_3MUTB, which neutralized BaL.26, ZM109F.PB4, or CAP45.2.00.G3 with $IC_{50}$ values of 24, 29, or 21 μg/mL, respectively (Fig. 3*A* and *SI Appendix*, Table S5).

The Abs we expressed were modeled to interact with HIV-1 due to a PG9-like HCDR3 structure. However, the crystal structure of PG9*wt* in complex with the CAP45.2.00.G3 V1/V2 scaffold revealed that there are additional contacts besides the HCDR3 that were involved in binding. We explored whether such contacts mediated the binding and neutralization observed in HIV-1–naïve donor-origin Abs, rather than the HCDR3 region. To test this, we expressed a variant Ab designated PG9 ΔCDR3 in which the HCDR3 was replaced with a GGG motif. This variant did not bind or neutralize any of the HIV strains tested. We also screened all variant Abs for epitope-specific neutralization by testing activity against the RHPA.N160A virus containing an N160A Env change that removes a glycosylation site required for PG9*wt* neutralizing activity (23). With the exception of one variant (VU30400_7MUT),

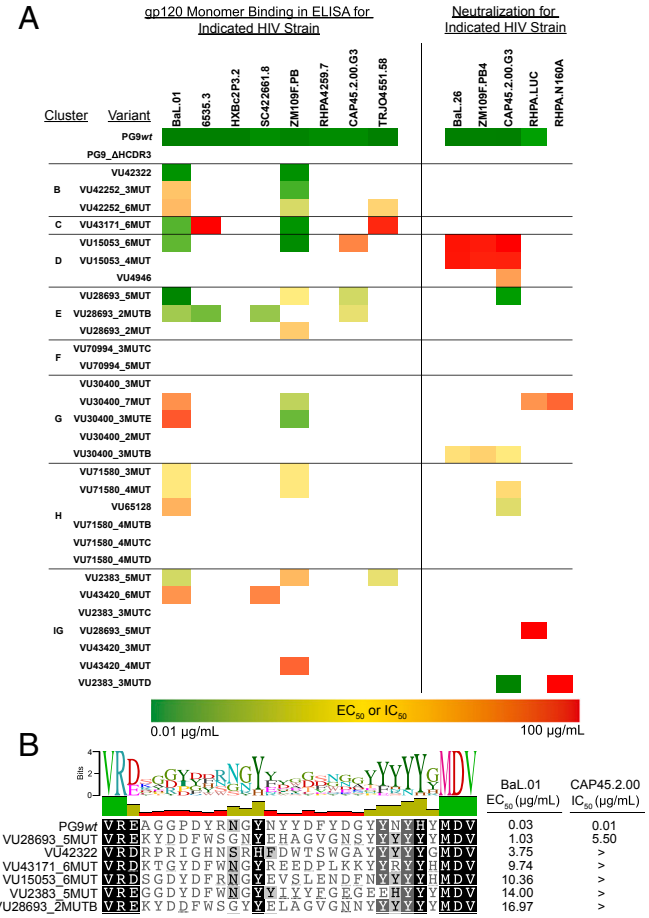


**Fig. 3.** Experimental characterization of HIV-1–naïve donor HCDR3 sequences that bind and neutralize HIV. (*A*) A blue–yellow–red heat map corresponding to the $EC_{50}$ or $IC_{50}$ for binding or neutralization, respectively. Values in white indicate an $EC_{50}$ or $IC_{50}$ higher than the top concentration of 100 μg/mL, and are shown for all characterized Ab variants that expressed at sufficient concentration for analysis. The viral variants used in binding and neutralization assays are shown at the top whereas the HCDR3 variants and parental Ab clusters are shown on the side. (*B*) The Ab variants that bound at a concentration less than 20 μg/mL are shown in a multiple sequence alignment. The black–gray scale indicates similarity to PG9. The $EC_{50}$ value is shown. All underlined amino acid sequences were redesigned from the parental Ab sequence.

the variants did not neutralize the N160A knockout mutant (Fig. 3*A* and *SI Appendix*, Table S5).

We explored the molecular basis for the functional and structural convergence of the clones on mimicry of PG9. Surprisingly, there was little sequence homology between the HCDR3s of HIV-1–naïve donor-reactive clones and that of PG9*wt* (Fig. 3*B*). With the exception of positions 93–95, $100_{C–E}$, and $100_{O}$–103, the sequence of each variant used a unique set of mutations that were predicted to achieve the hammerhead HCDR3 configuration and enable binding to HIV-1 gp120. We considered the success of the design process used here to introduce a minimal number of beneficial mutations into HCDR3 regions with suboptimal interaction, a process that mimics affinity maturation through somatic hypermutation. For instance, the Ab encoded by the *wt* HCDR3 clone for cluster E, termed VU11631, was found to express poorly in culture. In contrast, several minimally designed variant Abs, termed VU28693_2MUTB and VU28693_5MUT, were found to express, bind gp120, and neutralize HIV-1. VU28693_$Y100_{D}$N was predicted to pack more densely and form hydrogen bonds with K168 of the CAP45 V1/V2 interface, which mimics the interaction of PG9*wt*. VU28693_$L100_{G}$H exposes a polar group to the solvent face while retaining burial of the hydrophobic portion, and VU28693_$L100_{L}$N uses the terminal amine group to point toward solvent space and create a hydrogen bond with T163 of the V1/V2 interface (*SI Appendix*, Fig. S11*A*). VU15053_$W100_{B}$R is predicted to form two hydrogen bonds with the V1/V2 glycan at position N160, whereas VU15053_$R100_{L}$D forms an additional three hydrogen bonds that are out of the range of the *wt* arginine (*SI Appendix*, Fig. S11*B*).

## Discussion

A protective vaccine against HIV-1 likely will need to elicit a broadly neutralizing serum Ab response (24–29). Here we interrogated the long HCDR3 repertoire, before infection, using samples from a cohort of HIV-1–naïve donors, maximizing the sequence pool and diversity. The combination of computational modeling with bioinformatically driven heuristics allowed the screening of a large number of HIV-1–naïve Ab sequences. Although computational modeling of all 26,917 sequences might be feasible in principle, we sought to establish a method that could be extended to an orders-of-magnitude greater number of sequences that will result from the rapidly growing output of evolving amplicon sequencing technologies (30). We found that three out of the seven *wt* sequences we tested bound to HIV gp120, and each of these was derived from a unique donor. We chose the recombinant gp120 monomer to test the binding of these Abs. Note that some neutralizing Abs in this class fail to bind the gp120 recombinant monomer. This finding has been attributed to the fact that the epitope for Abs interacting with two protomers in the intact HIV BG505 trimer on the envelope is not fully recapitulated in a gp120 monomer (31). In addition, the gp120 monomer is differentially glycosylated from the native trimeric envelope (32). PG16, a long HCDR3-type Ab related to PG9, has been found to be more selective in the glycan types to which it binds and neutralizes (33). This finding is attributed to subtle mutations that interact with the branched sugars at positions 156 and 160.

Many Abs capable of adopting a PG9-like structure might be missed by a sequence-only search strategy. Instead, we sought to use structure-based homology prediction as a metric to assess predicted PG9-like HCDR3 activity. We attribute this finding to the many sequence-independent interactions found in PG9 and PG16 Ab complexes (11). In this way, a sequence may only need to adopt the rigid secondary structure found in the HCDR3 hammerhead structure.

We note the limitations of using only HCDR3 sequences rather than the full $V_H$ gene sequence or paired heavy–light chain variable gene sequences. We chose a method with shorter read length to achieve sequencing depth, as 30-amino acid-length HCDR3 sequences are infrequent in the human repertoire, and a large Ab sequence pool would be needed to find HCDR3 sequences with anything resembling predicted PG9-like HCDR3 activity. Indeed, we found two *wt* HCDR3 sequences with neutralizing activity, derived from 26,917 sequences that were potential candidates based on length from the HCDR3 repertoire. In the scope of this study, only 26,917 possible sequences were tested. We expect that the typical B-cell repertoire has more HCDR3 sequences containing a 30-amino acid-length HCDR3 than this representative sampling. With the exception of two Abs (VU30400_G_7MUT and VU2383_3MUTD), all of the predicted PG9-like HCDR3 sequences failed to neutralize RHPA.N160A, a PG9 epitope knockout mutation that removes the N160 glycan site. From this, we hypothesize that the HCDR3 sequences from the HIV-naïve repertoire work synergistically with the mature heavy- and light-chain mutations of PG9 by contributing to binding or neutralization.

Recently, Andrabi et al. reported work on viral variants that bind PG9 variant Abs with $V_H/V_L$ reverted mutations. Results in those studies show that the common $V_H$ germ-line genes in PG9 and CAP256.09 Abs play a vital role in the neutralization of sensitive viruses and that their ability to neutralize these viruses is not a function of HCDR3 alone. Instead, neutralization appeared to derive from interactions mediated by both the HCDR3 (especially the YYD amino acid motif from the common germ-line D gene)

and residues in $V_H$-encoded regions (including residues that are both germ line-encoded and somatic mutations). These investigators found that 16 of 26 mutations are necessary to bind a majority of the viral variants they tested, suggesting that some but not all of the somatic mutations outside of the HCDR3 of PG9 are required for recognition of HIV (34).

This work has several implications for HIV-1 vaccine design. The data reveal that some HIV-1–naïve donors possess B cells in the preimmune repertoire whose HCDR3 structure is predicted to mimic that of a class of V1/V2 binding Abs that is known to exhibit broad and potent HIV-1 neutralizing activity, when configured in the context of the naturally occurring PG9 antibody. Recently, a genetic pathway for the development of V1/V2 binding Abs with long HCDR3s has been elucidated for potent bnAbs (22), revealing a very large number and diversity of mutations from the unmutated common ancestor in the HCDR3 associated with neutralization breadth of diverse strains. That study suggests a large range of Ab sequence diversity in clonal lineages that can converge onto one epitope while maximizing breadth and potency. In the scope of this study, only 2 out of the 26,917 sequences from two different donors were found to have neutralizing activity against HIV. However, only 26,917 HCDR3 sequences of 30-amino acid length were identified in the deep-sequencing reactions used here, a much smaller number than the pool size of such antibodies that is estimated for the entire naïve B-cell repertoire in a person. Targeting the diverse repertoire of Abs with PG9-like HCDR3 structures and function in the preimmune repertoire may offer a more tractable focus for vaccination, because lower numbers of somatic mutations may be required for HIV-1 recognition by these clones. The data suggest that induction of B cells encoding long HCDR3s from the HIV-naïve repertoire is an attractive target for structure-based vaccine design.

## Materials and Methods

HCDR3 amplicons were generated from RNA that was extracted from the white blood cells of 70 HIV-uninfected individuals. Use of deidentified cells obtained from discarded leukofiltration filters was approved by the Vanderbilt and Red Cross Institutional Review Boards. Reads ($5.14 \times 10^7$) were obtained from using the Illumina HiSeq platform. Sequence tolerance for PG9 was modeled using the PG9 complex with scaffolded V1/V2 CAP45 (PDB ID code 3U4E) and the modeling suite Rosetta. Additional HCDR3 sequences were designed using Rosetta, starting with healthy donor HCDR3 sequences as template. Four HCDR3 sequences were synthesized and chimerized with the mature PG9 sequence. Chimeric Abs were tested for binding or neutralization using a gp120 ELISA or luciferase reporter TZM-bl assay, respectively. For full methods, see *SI Appendix, Materials and Methods*.


1. Ackerman M, Alter G (2013) Mapping the journey to an HIV vaccine. *N Engl J Med* 369(4):389–391.
2. Simek MD, et al. (2009) Human immunodeficiency virus type 1 elite neutralizers: Individuals with broad and potent neutralizing activity identified by using a high-throughput neutralization assay together with an analytical selection algorithm. *J Virol* 83(14):7337–7348.
3. Hraber P, et al. (2014) Prevalence of broadly neutralizing antibody responses during chronic HIV-1 infection. *AIDS* 28(2):163–169.
4. Mikell I, et al. (2011) Characteristics of the earliest cross-neutralizing antibody response to HIV-1. *PLoS Pathog* 7(1):e1001251.
5. Kwong PD, Mascola JR (2012) Human antibodies that neutralize HIV-1: Identification, structures, and B cell ontogenies. *Immunity* 37(3):412–425.
6. Corti D, Lanzavecchia A (2013) Broadly neutralizing antiviral antibodies. *Annu Rev Immunol* 31:705–742.
7. Klein F, et al. (2012) HIV therapy by a combination of broadly neutralizing antibodies in humanized mice. *Nature* 492(7427):118–122.
8. McCoy LE, Weiss RA (2013) Neutralizing antibodies to HIV-1 induced by immunization. *J Exp Med* 210(2):209–223.
9. Pejchal R, et al. (2010) Structure and function of broadly reactive antibody PG16 reveal an H3 subdomain that mediates potent neutralization of HIV-1. *Proc Natl Acad Sci USA* 107(25):11483–11488.
10. Pancera M, et al. (2010) 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* 84(16):8098–8110.
11. McLellan JS, et al. (2011) Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. *Nature* 480(7377):336–343.
12. Walker LM, et al.; Protocol G Principal Investigators (2009) Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. *Science* 326(5950):285–289.
13. Briney BS, Willis JR, Crowe JE, Jr (2012) Human peripheral blood antibodies with long HCDR3s are established primarily at original recombination using a limited subset of germline genes. *PLoS One* 7(5):e36750.
14. Leaver-Fay A, et al. (2011) ROSETTA3: An object-oriented software suite for the simulation and design of macromolecules. *Methods Enzymol* 487:545–574.
15. Rohl CA, Strauss CEM, Chivian D, Baker D (2004) Modeling structurally variable regions in homologous proteins with Rosetta. *Proteins* 55(3):656–677.
16. Berrondo M, Gray JJ (2011) Computed structures of point deletion mutants and their enzymatic activities. *Proteins* 79(10):2844–2860.
17. Larkin MA, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21):2947–2948.
18. Kuhlman B, Baker D (2000) Native protein sequences are close to optimal for their structures. *Proc Natl Acad Sci USA* 97(19):10383–10388.
19. Dal Porto JM, Haberman AM, Kelsoe G, Shlomchik MJ (2002) Very low affinity B cells form germinal centers, become memory B cells, and participate in secondary immune responses when higher affinity competition is reduced. *J Exp Med* 195(9):1215–1221.
20. Paus D, et al. (2006) Antigen recognition strength regulates the choice between extrafollicular plasma cell and germinal center B cell differentiation. *J Exp Med* 203(4):1081–1091.
21. Shih T-AY, Meffre E, Roederer M, Nussenzweig MC (2002) Role of BCR affinity in T cell dependent antibody responses in vivo. *Nat Immunol* 3(6):570–575.
22. Doria-Rose NA, et al.; NISC Comparative Sequencing Program (2014) Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies. *Nature* 509(7498):55–62.
23. Doores KJ, Burton DR (2010) Variable loop glycan dependency of the broad and potent HIV-1-neutralizing antibodies PG9 and PG16. *J Virol* 84(20):10510–10521.
24. Mascola JR, et al. (1999) Protection of macaques against pathogenic simian/human immunodeficiency virus 89.6PD by passive transfer of neutralizing antibodies. *J Virol* 73(5):4009–4018.
25. Burton DR, et al. (2012) A blueprint for HIV vaccine discovery. *Cell Host Microbe* 12(4):396–407.
26. Hessell AJ, et al. (2009) Broadly neutralizing human anti-HIV antibody 2G12 is effective in protection against mucosal SHIV challenge even at low serum neutralizing titers. *PLoS Pathog* 5(5):e1000433.
27. Hessell AJ, et al. (2010) Broadly neutralizing monoclonal antibodies 2F5 and 4E10 directed against the human immunodeficiency virus type 1 gp41 membrane-proximal external region protect against mucosal challenge by simian-human immunodeficiency virus SHIVBa-L. *J Virol* 84(3):1302–1313.
28. Hessell AJ, et al. (2009) Effective, low-titer antibody protection against low-dose repeated mucosal SHIV challenge in macaques. *Nat Med* 15(8):951–954.
29. Hessell AJ, Haigwood NL (2012) Neutralizing antibodies and control of HIV: Moves and countermoves. *Curr HIV/AIDS Rep* 9(1):64–72.
30. Finn JA, Crowe JE, Jr (2013) Impact of new sequencing technologies on studies of the human B cell repertoire. *Curr Opin Immunol* 25(5):613–618.
31. Julien J-P, et al. (2013) Asymmetric recognition of the HIV-1 trimer by broadly neutralizing antibody PG9. *Proc Natl Acad Sci USA* 110(11):4351–4356.
32. Doores KJ, et al. (2010) Envelope glycans of immunodeficiency virions are almost entirely oligomannose antigens. *Proc Natl Acad Sci USA* 107(31):13800–13805.
33. Pancera M, et al. (2013) Structural basis for diverse N-glycan recognition by HIV-1-neutralizing V1-V2-directed antibody PG16. *Nat Struct Mol Biol* 20(7):804–813.
34. Andrabi R, et al. (2015) Identification of common features in prototype broadly neutralizing antibodies to HIV envelope V2 apex to facilitate vaccine design. *Immunity* 43(5):959–973.