

# Supporting Information

West et al. 10.1073/pnas.1208984109

## SI Materials and Methods

### VRC01-Like Antibodies That Lack Some PVL Characteristic Residues.

Although many VRC01-like antibodies that lack one or more of the potent VRC01-like (PVL) characteristic residues are clearly less broad and potent than the main PVL antibodies (Abs), a few have nearly the same activity. For example, VRC-H5, VRC03 heavy chain paired with the VRC01 light chain, and 12A21 have very strong neutralization profiles (1–3). VRC03<sub>HC</sub>/VRC01<sub>LC</sub> and VRC-H5 lack Trp100<sub>B<sub>HC</sub></sub>. In the crystal structure of VRC03<sub>HC</sub>/VRC03<sub>LC</sub> complexed with gp120 (3), Asn279<sub>gp120</sub> is hydrogen bonded to Asp100<sub>C<sub>HC</sub></sub>, rather than Phe100<sub>D<sub>HC</sub></sub>, the residue that occupies the usual location for Trp100<sub>B<sub>HC</sub></sub>. Inspection of the sequence of VRC-H5 in this region suggests that the roles that Trp100<sub>B<sub>HC</sub></sub> normally plays can instead be performed by a large hydrophobic residue (e.g., Phe or Leu) at position 100B paired with a potential hydrogen bond donor or acceptor (e.g., Asn or Asp) at position 100A. Thus, the characteristic residues we have identified are not strictly necessary for potent neutralization. We suggest, however, that the most common path for Abs to develop broad VRC01-like activity is by possessing the PVL characteristic residues.

**Polymorphism at 50<sub>HC</sub> in the Human VH1-2 Gene.** The single-nucleotide polymorphism (SNP) that distinguishes VH1-2 alleles with Trp50<sub>HC</sub> (\*02, \*03, and \*04; nucleotide T) from those with Arg50<sub>HC</sub> (\*01 and \*05; nucleotide C) is RefSNP rs1065059 [National Center for Biotechnology Information dbSNP at [www.ncbi.nlm.nih.gov/projects/SNP/](http://www.ncbi.nlm.nih.gov/projects/SNP/)]. The minor allele frequency (those alleles encoding Arg50<sub>HC</sub>) is 23.0% (504 of 1,094 genomes or 2,188 chromosomes). Thus, if the minor allele is distributed evenly in the population, the fraction homozygous for the Arg50<sub>HC</sub> allele is  $0.23 \times 0.23 = 0.053 = 5.3\%$ .

**J Gene Segment Assignment for PVL Abs.** Previous analysis of the D and J<sub>H</sub> segments used by PVL Abs did not reveal a clear pattern of J gene segment use (1, 3). For example, VRC01, VRC-PG04, and VRC-CH31 Abs were reported to use IGHJ1\*01, IGHJ2\*01, and IGHJ4\*01, respectively (2, 3). However, the relatively high sequence identity between J<sub>H</sub> gene segments combined with the extensive hypermutation of PVL Abs makes the J<sub>H</sub> gene assignments uncertain.

One problem with definitive assignment of J<sub>H</sub> gene segments in Abs is that the number of nucleotides assumed to be removed by

exonuclease activity during V-D-J joining can affect the segment assigned to be the closest match. JoinSolver (4), which uses a different algorithm from that of ImmunoGeneTics (IMGT)/V-QUEST (5), assigns IGHJ1\*01 for VRC01 partially due to the assumption that 7 nt were removed from the J gene segment, whereas the average number of nucleotides removed from the 5' end of the J segment is 6 (6). Retention of the codon for Trp100<sub>B<sub>HC</sub></sub> within the J<sub>H</sub> gene segment requires removal of 4 or fewer nt. Examination of a set of human Abs derived from IGHJ2\*01 suggests that retention of these nucleotides is not uncommon (e.g., retention occurs in 7 of 10 of the IGHJ2\*01-containing Abs included in a previous study) (7).

To further explore the J gene segment assignment by IMGT/V-QUEST, we attempted to remove any bias caused by inclusion of the Trp100<sub>B<sub>HC</sub></sub>. Even if the Trp100<sub>B<sub>HC</sub></sub> codon is changed to CCT, a codon not found in any J<sub>H</sub> gene segment at that position, VRC-PG04 and 3BNC117 are still predicted by IMGT/V-QUEST to derive from IGHJ2\*01, whereas VRC01 and 12A12 are then predicted to derive from IGHJ1\*01 and IGHJ4\*02, respectively.

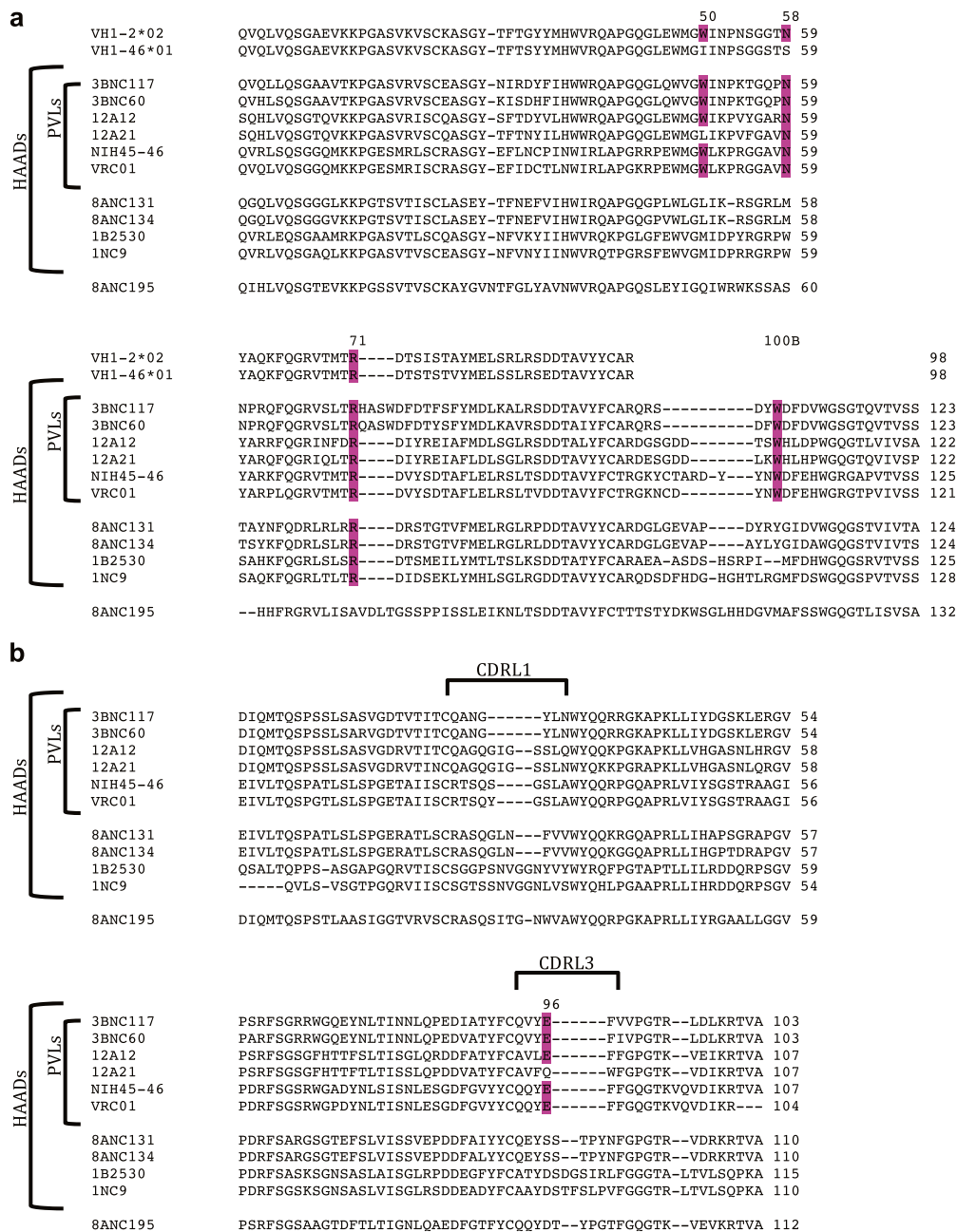
### Sequence Features That Vary Among PVL Abs, Including Insertions, Deletions, and Variations Within the CDRs.

Although PVL Abs share common features, other aspects of these Abs vary considerably. For example, the CDRH1 length varies between 10 in many PVLs; 11 in VRC-PG04, VRC-H12, VRC-H15; and 19 in VRC-CH31 and related clones (1–3). For PVLs VRC01 and NIH45–46, CDRH1 disulfide bonds with CDRH3. Although Trp100<sub>B<sub>HC</sub></sub> near the C-terminal end of CDRH3 is common to PVL Abs, the preceding portion of CDRH3 is highly variable. For some PVL Abs, this region of CDRH3 may contact the inner domain of gp120, as observed for NIH45–46 (8), whereas the shorter VRC01 CDRH3 does not.

In the light chain, the CDRL1 length is quite variable: 7 in the 3BNC clones; 9 in VRC01, NIH45–46, and VRC-PG04; and 11 in 12A12 and the VRC-CH31-related clones (1–3). The shorter versions are deletions from the germ line.

Several insertions also occur in PVL Abs. In VRC03, in framework region 3 (after Arg71<sub>HC</sub>) there is a seven-residue insertion. In 3BNC60 and related potent clonal variants, there is a four-residue insertion (sequence WDFD) in apparently the same location. When these Abs are complexed with gp120, these insertions are predicted to contact the V1V2 base in the bridging sheet (1).

1. Scheid JF, et al. (2011) Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding. *Science* 333:1633–1637.
2. Wu X, et al. (2010) Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science* 329:856–861.
3. Wu X, et al.; NISC Comparative Sequencing Program (2011) Focused evolution of HIV-1 neutralizing antibodies revealed by structures and deep sequencing. *Science* 333:1593–1602.
4. Souto-Carneiro MM, Longo NS, Russ DE, Sun HW, Lipsky PE (2004) Characterization of the human Ig heavy chain antigen binding complementarity determining region 3 using a newly developed software algorithm, JOINSOLVER. *J Immunol* 172:6790–6802.
5. Brochet X, Lefranc MP, Giudicelli V (2008) IMGT/V-QUEST: The highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. *Nucleic Acids Res* 36(Web Server issue):W503–W508.
6. Souto-Carneiro MM, Sims GP, Girschick H, Lee J, Lipsky PE (2005) Developmental changes in the human heavy chain CDR3. *J Immunol* 175:7425–7436.
7. Jackson KJ, Gaeta B, Sewell W, Collins AM (2004) Exonuclease activity and P nucleotide addition in the generation of the expressed immunoglobulin repertoire. *BMC Immunol* 5:19.
8. Diskin R, et al. (2011) Increasing the potency and breadth of an HIV antibody by using structure-based rational design. *Science* 334:1289–1293.



**Fig. S1.** Sequence alignment of CD4 binding site Abs. (A) Sequence alignment of highly active agonistic anti-CD4bs (HAAD) heavy chains, precursor germ-line V gene segments (Upper), and the 8ANC195 heavy chain (1). Characteristic PVL residues are highlighted in purple. (B) Sequence alignment of HAAD and 8ANC195 light chains (1). Characteristic PVL residues are highlighted in purple.



Full amino acid sequence:

(signal peptide)  
MGWSCIIILFLVATATGVHS

(VH)  
QVQLVQSGAEVKKPGASVKVSCASGYTFTGYYMHVWRQAPGQGLEWMGWINPNSG  
GTNYAQKFQGRVTMTRDSISTAYMELSRLRSDDTAVYYCARQRSDFWDFDLWGRGTLVTVSS

(CH1-CH3)  
ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSGLYSLSSVTVPSSSLGTQTYICNVNHKPSNTKVDKRVKVEPKSCDKTHTCPPCPAPELLG  
GPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQ  
YNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTKAKGQPREPQVYTLPPSR  
EEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLDSDGSFFLYSKLTVDK  
SRWQQGNVVFSCSVMEALHNHYTQKLSLSLSPGK

Alignment of 3BNC60 germline (GL) with 3BNC60 mature heavy chain:

```
GL      QVQLVQSGAEVKKPGASVKVSCASGYTFTGYYMHVWRQAPGQGLEWMGWINPNSGGTNYA 61
        QV L QSGA V KPGASV+VSC+ASGY + +++HW RQAPGQGL+W+GWINP +G N
3BNC60 QVHLSQSGAAVTKPGASVRSCEASGYKISDFIHWWRQAPGQLQVWVWINPKTGQPNNP 61

GL      QKFQGRVTMTR---DTSISTAYMELSRLSDDTAVYYCARQRSDFWDFLWGRGTLVTVSS 119
        ++FQGRV++TR D + YM+L +RSDDTA+Y+CARQRSDFWDFD+WG GT VTVSS
3BNC60 RQFQGRVSLTRQASWDFDTYSFYMDLKAVRSDDTAIYFCARQRSDFWDFVWVSGTQVTVSS 123
```

**Fig. S4.** Sequence of the 3BNC60 germ-line heavy chain used in the surface plasmon resonance binding experiment shown in Fig. 4. The underlined sections match VH1-2\*02 and IGHJ2\*01. The red section (mainly the D gene segment and added nucleotides) matches the mature 3BNC60 heavy chain. Point mutants were made at the residues shown in blue. The germ-line construct is also shown aligned with the mature 3BNC60 heavy chain.

```

NIH45-46      QVRLSQSGGQMKKPGESMRLSCRAS---GYE-----FLNCPINWIRLAPGRRPEWMCWL 51
VRC01        QVQLVQSGGQMKKPGESMRISCRAS---GYE-----FIDCTLNWIRLAPGKRPEWMCWL 51
12A12       SQHLVQSGTQVKKPGASVRISQAS---GYS-----FTDYVLHWRQAPGQGLEWMCWI 51
3BNC117     QVQLLQSGAAVTKPGASVRVSCAS---GYN-----IRDYFIHWRQAPGQGLQVWGI 51
3BNC60      QVHLSQSGAAVTKPGASVRVSCAS---GYK-----ISDHFHWRQAPGQGLQVWGI 51
VRC-CH31    QVQLVQSGAAVRKPGASVTVSKFAEDDDYSPYWVNPAPDFHIFLRLQAPGQGLEWLAWM 60
VRC-PG04    QVQLVQSGSGVKKPGASVRVSCWTS-EDIFE-----RTELHWRQAPGQGLEWICWV 52
IGHV1-2*02 # QVQLVQSGAEVKKPGASVKVSKAS---GYT-----FTGYMHWRQAPGQGLEWMCWI 51
IGHV1-46*01 QVQLVQSGAEVKKPGASVKVSKAS---GYT-----FTSYMHWRQAPGQGLEWMCWI 51
IGHV1-8*01  QVQLVQSGAEVKKPGASVKVSKAS---GYT-----FTSYDINWRQATGQGLEWMCWI 51
IGHV1-18*01 QVQLVQSGAEVKKPGASVKVSKAS---GYT-----FTSYGISWRQAPGQGLEWMCWI 51
IGHV1-45*01 # QMQLVQSGAEVKKTGSSVKVSKAS---GYT-----FTYRYLHWRQAPGQALEWMCWI 51
IGHV1-58*01 # QMQLVQSGPEVKKPGTSVKVSKAS---GFT-----FTSSAVQVWRQARGQRLEWICWI 51
IGHV1-3*01  QVQLVQSGAEVKKPGASVKVSKAS---GYT-----FTSYAMHWRQAPGQGLEWMCWI 51
IGHV1-69*01 QVQLVQSGAEVKKPGSSVKVSKAS---GGT-----FSSYAIHWRQAPGQGLEWMCWI 51
IGHV1-f*01  EVQLVQSGAEVKKPGATVKISCKVS---GYT-----FTDYMHWRQAPGKGLEWMCWI 51
IGHV1-24*01 QVQLVQSGAEVKKPGASVKVSKVS---GYT-----LTELSMHWRQAPGKGLEWMCWI 51
IGHV1-68*01 # QVQLGQSEAEVKKPGASVKVSKAS---GYT-----FTCCSLHWLQAPGQGLERMCWI 51
IGHV1-c*01  QVQLVQSWAEVRKSGASVKVSCSFS---GFT-----ITSYGIHWRQSPGQGLEWMCWI 51
                                                    50

NIH45-46      KPRGGAVNY-ARKFQGRVTMTR----DVYSDTAFLELRSLTSDDTAVYFCTRGRKYCTARD 106
VRC01        KPRGGAVNY-ARPLQGRVTMTR----DVYSDTAFLELRSLTVDDTAVYFCTRGNKCD--- 103
12A12       KPVIYGARNY-ARRFQGRINFDK----DIYREIAFMDLSGLRSDDTALYFCARDGSGDD-- 104
3BNC117     NPKTGQPNN-PRQFQGRVSLTRHASWDFDTFSFYMDLKALRSDDTAVYFCAR-----QR 104
3BNC60      NPKTGQPNN-PRQFQGRVSLTRQASWDFDTFSFYMDLKAVRSDDTAVYFCAR-----QR 104
VRC-CH31    NPTNGAVNY-AWYLNQGRVTATR----DRSMTAFLEVKSLRSDDTAVYCARAQKRRGR-- 113
VRC-PG04    KTVTGAVNYFGSPDFRQRVSLTR----DRDLFTAHDIRGLTQGDATYFCARQKFFYTG-- 107
IGHV1-2*02 # NPNSGGTNY-AQKFQGRVTMTR----DTSISTAYMELSLRSDDTAVYYCAR----- 98
IGHV1-46*01 NPSGGSTY-AQKFQGRVTMTR----DTSTSTVYMESSLRSEDATVYYCAR----- 98
IGHV1-8*01  NPNSGNTGY-AQKFQGRVTMTR----NTSISTAYMELSSLRSEDATVYYCAR----- 98
IGHV1-18*01 SAYNGNTNY-AQKLQGRVTMTT----DTSTSTAYMELSLRSDDTAVYYCAR----- 98
IGHV1-45*01 # TPFNGNTNY-AQKFQDRVTITR----DRSMSTAYMELSSLRSEDATVYYCAR----- 98
IGHV1-38*01 # VVGSNTNY-AQKFQERVITR----DMSTSTAYMELSSLRSEDATVYYCAR----- 98
IGHV1-3*01  NAGNGNTKY-SQKFQGRVTITR----DTSASTAYMELSSLRSEDATVYYCAR----- 98
IGHV1-69*01 IPIFGTANY-AQKFQGRVTITA----DESTSTAYMELSSLRSEDATVYYCAR----- 98
IGHV1-f*01  DPEDGETIY-AEKFQGRVTITA----DTSTDTAYMELSSLRSEDATVYYCAT----- 98
IGHV1-24*01 DPEDGETIY-AQKFQGRVTMTE----DTSTDTAYMELSSLRSEDATVYYCAT----- 98
IGHV1-68*01 # TLYNGNTNY-AKKFQGRVTITR----DMSLRTAYIELSSLRSEDSAVYYWAR----- 98
IGHV1-c*01  NPNGNSPSY-AKKFQGRFTMTR----DMSTTAYITDLSSLTSEDMAVYYAR----- 98
                                                    58          71

NIH45-46      YYNWDFEHWGRGAPVTVSS 125
VRC01        -YNWDFEHWGRGTPVIVSS 121
12A12       -TSWHLDPWGQGLVIVSA 122
3BNC117     SDYWDFDVGSGTQVTVSS 123
3BNC60      SDFWDFDVGSGTQVTVSS 123
VRC-CH31    -SEWAYAHWGQGTPTVVVSS 131
VRC-PG04    -QQWYFDLWGRGTLVIVSS 125
                                                    100B

```

Fig. S5. Alignment of PVL heavy-chain sequences with IGHV1 gene segments. PVL characteristic residues are highlighted in purple. Germ-line segments indicated with “#” have PVL characteristic residues Trp50<sub>HC</sub>, Asn58<sub>HC</sub>, and Arg71<sub>HC</sub>.

VRC01	Clone	270	280	450	460	470
IC50	Consensus	EVVIRSSNFTDIAKIIII	IRGLLLTRDGG	-----	NSTNNTETFRPG	
	HxBc2	EVVIRSSNFTDIAKIIII	IRGLLLTRDGG	-----	NSNNESEIFRPG	
0.017	1_95C5	DI...E...K...T...N...	.....V.....	.....	-----ENVV	
0.020	45_01dG5	I...E...IK.....	.....	.....	-----SSTNGT	
0.041	1_01B5	DI...E...K...TNT..	.....	.....	-----EDQ...I...	
0.049	45_01dH5	.....D...K.....	.....	.....	-----G.....	
0.089	N26_07TC40	..I...D...SN...T...	.....	.....	-----NEDNGGP.....	
0.143	N90_08B11	.....	.....	.....	-----IN.....	
0.328	N90_08B16	.....	.....	.....	-----KN.....	A
0.372	N90_08A19	.....	.....	.....	-----KN.....	A
0.591	N26_07A38	..I...A...LM...T...	..I.....	.....	-----D.KDN.....	A
0.644	N90_08B6	.....	.....	.....	-----KN.....	A
0.661	18_99A27	..I...E...I...N...T...	.....	.....	-----NETNKT...I...	A
0.703	N90_08B2	.....	.....	.....	-----IN.....	A
0.855	18_99B19	..I...E...ISN...T...	.....	.....	-----HGN.ETD.....	E
1.000	N90_08A15	.....	.....	.....	-----IN.....	A
1.100	18_99TB13	..I...E...I...N...T...	.....	.....	-----HGN.ETD.....	E
1.200	18_99B51	..I...E...ISN...T...	.....	.....	-----HGN.ETD.....	E
1.300	N90_08A16	.....	.....	.....	-----IN.....	A
1.300	N90_08A25	.....	.....	.....	-----KN.....	A
1.600	N26_07A21	..I...A...S...T...	.....	.....	-----NNNDRTNNGSDS..NSTQ..I...	A
1.600	18_99B46	..II...E...I...N...T...	.....	.....	-----HGN.ETD.....	E
1.700	N90_08B17	.....	.....	.....	-----IN..N.....	A
1.900	18_99B15	..I...E...L...S...V...T...	.....	.....	-----HGN.ETD.....	E
2.300	B7B5_88A2	.....A...SN...V...T...	.....	.....	-----E..Q.....	
2.500	N90_08A13	.....	.....	.....	-----IN..N.....	A
2.900	N90_08B4	.....	.....	.....	-----IN..N.....	A
3.100	B7B5_88A3	.....A...SN...V...T...	.....	.....	-----E..QN.....	
3.500	B7B5_88A6	.....A...SN...V...T...	.....	.....	-----E..Q.....	
3.500	18_99A34	..I...E...L...S...V...T...	.....	.....	-----HGN.ETD.....	E
3.600	45_01dH1	..II...E...L...T...	.....	.....	-----NI..TNETT.....	
4.800	18_99B1	..I...E...I...N...T...	.....	.....	-----HGN.ETD.....	E
9.600	B7B5_88TB2	.....E...S...V...N...	.....	.....	-----YK.....NESN.SNESKP.....	
11.400	B7B5_88TB3	.....E...S...V...N...	.....	.....	-----YK.....NESN.SNESKP.....	
11.500	1_01A20	DI..L...E...K...SNN..	.....	.....	-----NT.SQS.T.....	
14.600	1_95TC14	DI...E...K...T...T...	.....	.....	-----KT..QS.T.....	
17.000	B7B5_88A1	.....A...S...V...T...	.....	.....	-----GN.E..P.N.....	
17.500	45_06A3	..IM...PK...ISE...DD...	.....	.....	-----NIT..E.IT.....Q	
19.000	N26_07B18	..II...EDL...R...E...T...	.....	.....	-----NSST..NNGSDS..NSTK..I...	
19.300	N26_07TC39	..I...A...M...T...	.....	.....	-----NDTNHEN.....	
25.900	B7B5_88A10	.....A...S...V...T...	.....	.....	-----V.....-ESSSN.....	
31.200	45_06C7	..IM...PK...ISE...DD...	.....	.....	-----NIT..E.IT.....Q	
50.000	45_01A8	..II...E...L...K...Q...T...	.....	.....	-----NI..TNE.T.....	
50.000	45_01A10	..II...E...L...E...H...T...	.....	.....	-----NI..TNK.T.....E	
50.000	45_01A14	..II...R...LSE...DDT...	.....	.....	-----NI..TNETT.....	
50.000	45_01A21	..II...E...L...K...Q...T...	.....	.....	-----NI..TNK.T.....E	
50.000	45_01B14	..II...E...L...R...Q...T...	.....	.....	-----NI..TNK.T.....E	
50.000	45_01E8	..II...E...L...E...H...T...	.....	.....	-----NI..TNE.T.....	
50.000	45_01E11	..II...E...L...E...H...T...	.....	.....	-----NI..TNK.T.....E	
50.000	45_01E17	..II...E...L...E...H...T...	.....	.....	-----NI..TNK.T.....E	
50.000	45_01F4	..II...E...L...E...H...T...	.....	.....	-----NI..TNE.T.....	
50.000	45_06A1	..IM...PK...LSE...DD...	.....	.....	-----NI..TDK.....Q	
50.000	45_06A2	..II...E...L...A...H...T...	.....	.....	-----NIT..E.IT.....Q	
50.000	45_06A4	..II...E...L...E...H...T...	.....	.....	-----NI..TDK.....Q	
50.000	45_06A8	..II...E...L...A...H...T...	.....	.....	-----NIT..E.IT.....Q	
50.000	45_06A10	..I...R...LSE...DD...	.....	.....	-----NIT..E.IT.....Q	
50.000	45_06C8	..II...E...L...A...H...T...	.....	.....	-----NIT..E.IT.....E	
50.000	45_09A5	..II...E...L...E...H...T...	.....	.....	-----NIT..K..I.T.....Q	
50.000	45_09A6	..I...R...LSE...D...	.....	.....	-----NIT..E.IT.....Q	
50.000	45_09A10	..IM...PK...ISE...DD...	.....	.....	-----NIT..E.IT.....Q	
50.000	45_09B3	..I...R...LSE...D...	.....	.....	-----NIT..E.IT.....Q	
50.000	45_09C5	..I...E...L...A...H...T...	.....	.....	-----D.....IT..E.IT.....Q	
50.000	45_09C8	..IM...PK...ISE...DD...	.....	.....	-----NIT..E.IT.....Q	
50.000	45_09C20	..II...E...L...E...H...T...	.....	.....	-----NIT..E.IT.....Q	
50.000	45_09C26	..IM...PK...ISE...DD...	.....	.....	-----NIT..E.IT.....Q	
50.000	45_09TA1	..I...R...LSE...D...	.....	.....	-----NIT..E.IT.....Q	
50.000	1_95A21	DI...E...K...T...T...	.....	.....	-----ENVV..I.....	
50.000	1_95C1	DI...E...K...T...T...	.....	.....	-----ENVV.....	
50.000	1_95C10	DI...E...K...T...T...	.....	.....	-----ENVV.....	
50.000	1_95TC1	DI...E...H...T...T...	.....	.....	-----EDQ..K..I.....	
50.000	1_95TC6	DI...E...K...T...T...	.....	.....	-----ENVV..I.....	
50.000	1_95TC13	DI...E...H...T...T...	.....	.....	-----EDO..K.....	
50.000	1_01A10	DI..L...E...K...SNN..	.....	.....	-----NT.SQS.T.....	
50.000	1_01A19	DI..L...E...K...SNN..	.....	.....	-----NT.SQS.M.....	
50.000	1_01B3	DI..L...E...E...SNN..	.....	.....	-----NT.SQS.M.....	
50.000	1_01TB5	DI..L...E...E...S...T...	.....	.....	-----NT.SQS.T.....	
50.000	1_06B1	DI..L...E...K...SNN..	.....	.....	-----NT..QS.T.....	
50.000	1_06B3	DI..L...E...E...SNN..	.....	.....	-----NT..QS.A.....	
50.000	1_06B7	DI..L...E...K...SNN..	.....	.....	-----NT..QS.T.....	
50.000	1_06B9	DI..L...E...K...SNN..	.....	.....	-----NT..QS.T.....	
50.000	1_06C3	DI..L...E...E...SNN..	.....	.....	-----NT..QS.T.....	
50.000	1_06C5	DI..L...E...E...SNN..	.....	.....	-----NT.SQS.T.....	
50.000	1_06D29	DI..L...E...K...SNN..	.....	.....	-----NT..QS.T.....	
50.000	N26_07A10	..II...EDLN...H...N...	.....	.....	-----D.KDN.....	
50.000	N26_07A14	..II...EDL...R...E...T...	.....	.....	-----D.KDN.....	
50.000	N26_07A16	..I...EDLN...H...N...	.....	.....	-----N..N.....	
50.000	N26_07A22	..I...EDLN...H...N...	.....	.....	-----NSST..NNGSDS..NSTK..I...	
50.000	N26_07B9	..II...EDL...R...E...T...	.....	.....	-----NETKHEN.....	
50.000	N26_07B10	..II...EDL...K...E...T...	.....	.....	-----D..DN.....	
50.000	N26_07B22	..II...EDLM...K...E...T...	.....	.....	-----NNNSSTNNGSDS..NSTK..I...	
50.000	N26_07B36	..I...A...LM...T...	.....	.....	-----NETKHEN.....	
50.000	N26_07B41	..I...EDLN...H...N...	.....	.....	-----D.KDN.....	
50.000	B7B5_88A4	.....E...S...V...T...	.....	.....	-----V.....-QNSSN.....	
50.000	B7B5_88A5	.....E...S...V...T...	.....	.....	-----DN.E..SIN.....	
50.000	B7B5_88A7	.....E...S...E...V...T...	.....	.....	-----D..SIN.....	
50.000	B7B5_88TB1	.....E...S...V...T...	.....	.....	-----V.....-QNSSN.....	
50.000	B7B5_88TB6	.....E...S...V...T...	.....	.....	-----V.....-ENSSD.....	
50.000	18_99A38	..I...E...L...S...V...T...	.....	.....	-----YGN.ETD.T.....	
50.000	18_99B27	..I...E...L...S...V...T...	.....	.....	-----HGN.ETD.....E	
50.000	18_99B39	..I...E...I...N...T...	.....	.....	-----NETNKT.....E	
50.000	18_99B48	..I...E...ISN...T...	.....	.....	-----NETNKT.....E	
50.000	18_99TA3	..I...E...S...V...T...	.....	.....	-----HGN.DTD.T.....	

**Fig. S6.** Sequences of envelopes from donor 45 and four other individuals as reported in ref. 1. Sequences for residues 269–285 and 449–471 are shown; strains are listed in order of their VRC01 IC<sub>50</sub>. Residues 279, 280, and 456–459 are highlighted.

1. Wu X, et al. (2012) Selection pressure on HIV-1 envelope by broadly neutralizing antibodies to the conserved CD4-binding site. *J Virol* 86:5844–5856.

**Table S1. Classification of CD4-binding site Abs**

Class	Members
PVL Abs, neutralizing >75% of strains	VRC01, VRC02 NIH45–46 3BNC60, 3BNC117, 3BNC62, 3BNC95, 3BNC176 12A12 VRC-PG04 VRC-CH31, VRC-CH30, VRC-CH32, VRC-CH33, VRC-CH34 VRC03HC/VRC01LC gVRC-H5(d74)/VRC-PG04LC gVRC-H12(d74)/VRC-PG04LC*
Almost PVL Abs, related to known PVL Abs, neutralizing 10–75% of strains	VRC03 VRC01HC/VRC03LC 3BNC55, 3BNC91, 3BNC104, 3BNC89 12A21 VRC-PG04b
Defective PVL Abs, related to known PVL Abs, neutralizing <10% of strains	3BNC42, 3BNC53, 3BNC66, 3BNC72, 3BNC102 3BNC108, 3BNC142, 3BNC153, 3BNC156, 3BNC158 Germ-line parents of PVL Abs
Non-PVL HAADs	8ANC131 8ANC134 1B2530 1NC9 And related clones
Other CD4-binding site Abs	HJ16 <sup>†</sup> b12 b13 F105 8ANC195

Antibodies are described in refs. 1–4.

\*Plus seven additional HC clones (paired with PG04 LC) from deep sequencing of donor 74.

<sup>†</sup>HJ16 sequences (patent WO2011092593) do not appear to be particularly similar to PVLs. CDRL3 is 9 aa long, the heavy chain does not derive from VH1-2, and the Ab is missing several of the signature PVL residues.

1. Corti D, et al. (2010) Analysis of memory B cell responses and isolation of novel monoclonal antibodies with neutralizing breadth from HIV-1-infected individuals. *PLoS ONE* 5:e8805.
2. Scheid JF, et al. (2011) Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding. *Science* 333:1633–1637.
3. Wu X, et al. (2010) Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science* 329:856–861.
4. Wu X, et al.; NISC Comparative Sequencing Program (2011) Focused evolution of HIV-1 neutralizing antibodies revealed by structures and deep sequencing. *Science* 333:1593–1602.

**Table S2. Cross-species analysis for the presence of PVL characteristic residues in heavy-chain V gene segments**

Species	No. V <sub>H</sub> s examined	No. w/W50 and R71	No. w/W50, N58, and R71
Human	70	7*	4 <sup>†</sup>
Mouse	269	0	0
Rat	204	1 <sup>‡</sup>	0
Rabbit	49	2 <sup>§</sup>	0
Sheep	10	0	0
Pig	5	0	0
Rhesus macaque	89	1 <sup>¶</sup>	0

V<sub>H</sub> gene segment sequences from the IMGT database were examined, except for the macaque sequences, which were obtained from the Rhesus macaque Ig gene database ([http://www.kcl.ac.uk/immunobiology/Mac\\_ig/](http://www.kcl.ac.uk/immunobiology/Mac_ig/)). For the human genes, all V<sub>H</sub> alleles in IMGT were also checked.

\*VH1-2\*02, VH1-45\*01, VH1-58\*01, VH1-68\*01, VH1-c\*01, VH1-3\*01, and VH1-8\*01.

<sup>†</sup>VH1-2\*02, VH1-45\*01, VH1-58\*01, and VH1-68\*01 (pseudogene).

<sup>‡</sup>IGHV12S1\*01.

<sup>§</sup>IGHV1555\*01 and IGHV1556\*01.

<sup>¶</sup>IGHV11.

**Table S3. HIV-1 strains with sequence variations at critical PVL contact residues: Analysis based on NIH45–46 data as described in ref. 1**

Strain	Critical contact residues on gp120					IC <sub>50</sub> , μg/mL			
	279	280	456	458	459	VRC01	NIH45–46	VRC-PG04	3BNC117
Residue in most PVL-sensitive strains	Asx	Asn	Arg	Gly	Gly				
620345_c1	Lys		Ser		Asp	>50	>50	>50	>15
89_F1_2_25			Ser	Asn		ND	>50	ND	>50
6540_v4_c1		Ser	Ser	Tyr		>50	>50	>50	>50
6545_v4_c1		Ser	Ser	Tyr		>50	>50	ND	>50
Du422.1			Trp			>50	>50	>50	>50
T250_4					Pro	>50	>50	>50	>15
T278_50	Ala				Glu	>50	>50	>50	>15
Ce1172_H1					–/Ile	ND	>50	ND	>50
X2088_c9					Val	>50	>50	>50	>50
H086.8	Lys				Asp	>50	>30	>50	>15

Neutralization data are from refs. 2–4. Position 279<sub>gp120</sub> was previously shown to be important in resistance to VRC01; specifically the effects of mutation D279A<sub>gp120</sub> in JRC5F, T278\_50, BL01, and TV1\_29 viral strains were evaluated (5).

1. Diskin R, et al. (2011) Increasing the potency and breadth of an HIV antibody by using structure-based rational design. *Science* 334:1289–1293.
2. Scheid JF, et al. (2011) Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding. *Science* 333:1633–1637.
3. Wu X, et al. (2010) Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science* 329:856–861.
4. Wu X, et al. (2009) Mechanism of human immunodeficiency virus type 1 resistance to monoclonal antibody B12 that effectively targets the site of CD4 attachment. *J Virol* 83:10892–10907.
5. Li Y, et al. (2011) Mechanism of neutralization by the broadly neutralizing HIV-1 monoclonal antibody VRC01. *J Virol* 85:8954–8967.