Supporting Information

Kwon et al.

SI Materials and Methods

Design of pVRC8400-HIV-1 clade B_{YU2} ΔV123 expression construct. The pVRC8400-HIV-1 clade B $_{YU2}$ 44-492ΔV123 ($core_e$) expression construct was made by inserting a leader sequence encoding mouse IL-2 MYSMQLASCVTLTLVLLVN followed by the codon-optimized sequence encoding …
^{44}VWKEATTTLFCASDAKAYDTEVHNVWATHACVPTDPNPQEVKLENVTENFNMWKNNMV
EQMHEDIISLWDQSLKPCVKLTGGSVITQACPKVSFEPIPIHYCAPAGFAILKCNDKKF
NGTGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAEEEIVIRSENFTNNAKTIIVQLNESV
VINCTRPNNGGSGSGGDIRQAHCNLSKTQWENTLEQIAIKLKEQFGNNKTIIFNPSGGD
PEIVTHSFNCGGEFFYCNSTQLFTWNDTRKLNNTGRNITLPCRIKQIINMWQEVGKAMY
APPIRGQIRCSSNITGLLLTRDGGKDTNGTEIFRPGGGDMRDNWRSELYKYKVVKIE^{492}
… at the 5' XbaI and 3' BamHI sites. The insert was synthesized (GENEART) to replace the sequence encoding ...
^{124}PLCVTLNCTDLRNATNTTSSSWETMEKGEIKNCSFNITTSIRDKVQKEYALFYNLDV
VPIDNASYRLISCNT^{198} ... from the V1/V2 loop with a sequence encoding a GG and to replace the sequence encoding ... ^{302}NTRKSINIGPGRALYTTGEII^{323} ... from the V3 loop with a sequence encoding a GGSGSG linker.

Design of pVRC8400-HIV-1 clade C_{1086} ΔV123 expression construct. The codon-optimized pVRC8400-HIV-1 clade C_{1086} 44-492 ΔV123 ($core_e$) expression construct was made by inserting a leader sequence encoding mouse IL-2 MYSMQLASCVTLTLVLLVN followed by the sequence encoding ...
^{44}VWKEAKTTLFCASDAKAYEKEVHNVWATHACVPTDPNPQEMVLANVTENFNMWKNDMV
EQMHEDIISLWDESLKPCVKLTGGSAITQACPKVSFDPIPLHYCAPAGFAILKCNNKTF
NGTGPCRNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTNNAKTIIVHLNESV
NIVCTRPNNGGSGSGGNIRQAHCNINESKWNNTLQKVGEELAKHFPSKTIKFEPSSGGD
LEITTHSFNCRGEFFYCNTSDLFNGTYRNGTYNHTGRSSNGTITLQCKIKQIINMWQEV
GRAIYAPPIEGEITCNSNITGLLLLRDGGQSNETNDTETFRPGGGDMRDNWRSELYKYK
VVEIK^{492} … at the 5' XbaI and 3' BamHI sites. The insert was synthesized (GENEART) to replace the sequence encoding ...
^{124}PLCVTLNCTNVKGNESDTSEVMKNCSFKATTELKDKKHKVHALFYKLDVVPLNGNSS
SSGEYRLINCNT^{198} ... from the V1/V2 loop with a sequence encoding a GG and to replace the sequence encoding ... ^{302}NTRKSIRIGPGQTFYATGDII^{323} ... from the V3 loop with a sequence encoding a GGSGSG linker.

Design of pVRC8400-HIV-1 clade A/E $_{93TH057}$ ΔV123 expression construct. The codon-optimized pVRC8400-HIV-1 clade A/E $_{93TH057}$ 44-492 ΔV123 ($core_e$) expression construct was made by inserting a leader sequence encoding mouse interleukin-2 (IL-2) MYSMQLASCVTLTLVLLVN followed by the sequence encoding …
^{44}VWKDADTTLFCASDAKAHETEVHNVWATHACVPTDPNPQEIHLENVTENFNMWKNNMV
EQMQEDVISLWDQSLQPCVKLTGGSVIKQACPKISFDPIPIHYCTPAGYVILKCNDKNF
NGTGPCKNVSSVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTNNAKTIIVHLNKSV

EINCTRPSNGGSGSGGDIRKAYCEINGTKWNKVLKQVTEKLKEHFNNKTIIFQPPSGGD
LEITMHHFNCRGEFFYCNTTQLFNNTCIGNETMKGCNGTITLPCKIKQIINMWQGTGQA
MYAPPIDGKINCVSNITGILLTRDGGANNTSNETFRPGGGNIKDNWRSELYKYKVVQIE[4
92] … at the 5' XbaI and 3' BamHI sites. The insert was synthesized (GENEART) to replace the sequence encoding ...
[124]PLCVTLHCTTAKLTNVTNITNVPNIGNITDEVRNCSFNMTTEIRDKKQKVHALFYKL
DIVQIEDKNDSSKYRLINCNT[198] … from the V1/V2 loop with a sequence encoding a GG and to replace the sequence encoding ...
[302]NMRTSMRIGPGQVFYRTGSIT[323] ... from the V3 loop with a sequence encoding a GGSGSG linker.

Design of pVRC8400-HIV-1 clade $C_{ZM109F.PB4}$ ΔV123 expression construct. The codon-optimized pVRC8400-HIV-1 clade $C_{ZM109F.PB4}$ 44-492 ΔV123 ($core_e$) expression construct was made by inserting a leader sequence encoding mouse interleukin-2 (IL-2) MYSMQLASCVTLTLVLLVN followed by the sequence encoding …
[44]VWKEAKTTLFCASDAKSYEREVHNVWATHACVPTDPDPQELVMANVTENFNMWKNDMV
DQMHEDIISLWDQSLKPCVKLTGGSTITQACPKVSFDPIPIHYCAPAGYAILKCNNKTF
SGKGPCSNVSTVQCTHGIRPVVSTQLLLNGSLAEEEIVIRSENLTDNAKTIIVHLNKSV
EIECIRPGNGGSGSGGDIRKAYCKINGSEWNETLTKVSEKLKEYFNKTIRFAQHSGGDL
EVTTHSFNCRGEFFYCNTSELFNSNATESNITLPCRIKQIINMWQGVGRAMYAPPIRGE
IKCTSNITGLLLTRDGGNNNNSTEEIFRPEGGNMRDNWRSELYKYKVVEIK[492] … at the 5' XbaI and 3' BamHI sites. The insert was synthesized (GENEART) to replace the sequence encoding ...
[124]PLCVTLNCTSPAAHNESETRVKHCSFNITTDVKDRKQKVNATFYDLDIVPLSSSDNS
SNSSLYRLISCNT[198] ... from the V1/V2 loop with a sequence encoding a GG and to replace the sequence encoding ... [302]NTRKSIRLGPGQTFYATGDVI[323] ... from the V3 loop with a sequence encoding a GGSGSG linker.

Expression, purification, and crystallization. Detailed procedures of purification and deglycosylation of gp120 glycoproteins were reported previously (1). Briefly, supernatants of transiently transfected 293F or HEK293S $GnTi^-$ cell culture medium expressing gp120 variants were passed through the 17b conjugated- protein A column, washed with 1X PBS, and eluted with IgG Elution buffer (Pierce). 17b column-purified gp120s were deglycosylated with Endo H and further purified with a Concanavalin A (Sigma) column to remove glycosylated gp120. The final step of purification was to use a Superdex 200 column (GE Healthcare) to obtain homogeneous proteins. Proteins were concentrated to ~10 mg/ml in 2.5 mM Tris-HCl, pH 7.5, 350 mM NaCl, 0.02% NaN_3. Crystals grew in (16% PEG 3350, 0.1 M $CaCl_2$, 0.1M Tris-HCl pH 7.5), (16% PEG 1500, 0.1M $CaCl_2$, 0.1M imidazole pH 6.5), (10% PEG 3350, 10% iso-propanol, 0.2 M Ammonium citrate pH 7.5), and (14% PEG 1500, 12% PEG 400, 0.1M HEPES pH 7.5) using hanging-drop vapor diffusion method at 293 K for unliganded clade B_{YU2}, clade C_{1086}, clade $C_{ZM109F.PB4}$, and clade $E_{93TH057}$, respectively. Co-crystals of NBD-556 with clade C_{1086} grew in the same conditions that produced unliganded clade C_{1086} crystals. For co-crystallization, proteins were prepared to have 100 μM of NBD-556 in 2.5 mM Tris-HCl, pH 7.5, 350 mM NaCl, 0.02% NaN_3, and 5% DMSO.

Data collection, structure determination, and refinement. A single crystal was flash-frozen using a liquid nitrogen stream in cryo-protection solution. Cryo-protection solutions were prepared to contain (7.5% 2r3r-Butanediol, 18% PEG 3350, 0.1M $CaCl_2$, 0.1M Tris-HCl 7.5), (20% Glycerol, 18% PEG 1500, 0.1M $CaCl_2$, 0.1M imidazole pH 6.5), (15% 2r3r-Butanediol, 12% PEG 3350, 10% iso-propanol, 0.2M Ammonium citrate pH 7.5), and (30% Ethylene glycol, 16% PEG 1500, 12% PEG 400, 0.1M HEPES pH 7.5) for the unliganded clade B_{YU2}, clade C_{1086}, clade $C_{ZM109F.PB4}$, and clade $E_{93TH057}$, respectively. Data sets were collected to 3.4 Å, 2.8 Å, 4.0 Å and 1.9 Å resolution for clade B_{YU2}, clade C_{1086}, clade $C_{ZM109F.PB4}$, and clade $E_{93TH057}$, respectively, at the Advanced Photon Source (beamline SER-CAT ID22), indexed, integrated, and scaled with the HKL2000 package (2). Data collected from a crystal of clade B_{YU2} showed twinning (twin fraction: 0.29, twin operator=h, -h-k, -l). Initial trial of molecular replacement with the program Phaser (3) using unliganded SIV $core_{min}$ (PDB ID: 2BF1) as search model did not produce any solution. Therefore, we proceeded with the CD4-bound gp120 $core_e$ (PDB ID: 3JWD) as the search model using PHENIX (4) and found three correct molecules out of four in the asymmetric unit. [RFZ=4.9 TFZ=15.0 PAK=0 LLG=199 RFZ=3.4 TFZ=24.2 PAK=0 LLG=579 RFZ=3.2 TFZ=21.1 PAK=0 LLG=779]. To find the remaining molecule, we rigid-body refined the three found molecules, and generated a *2Fo-Fc* map using only the outer domains (256-474), without part of the bridging sheet (422-440), of the three molecules to minimize model bias. Electron density of the 2Fo-Fc map unambiguously showed that the inner domain and bridging sheet conformations were the same as those found in the CD4-bound conformation. Then we generated another 2Fo-Fc map using the entire CD4-bound gp120 core. At this point, the electron density of the remaining molecule in the asymmetric unit was clear enough that we were able to fit it manually using Coot (5). Molecular replacement of the data collected from unliganded clade C_{1086}, $E_{93TH057,}$ and $C_{ZM109F.PB4}$ were performed with one (chain A) of four molecules in the unliganded clade B_{YU2} gp120. Using PHENIX AutoMR, we found two [RFZ=12.8 TFZ=23.5 PAK=4 LLG=566 RFZ=9.5 TFZ=33.3 PAK=4 LLG=1345 LLG=1368], one molecule [RFZ=12.3 TFZ=21.4 PAK=0 LLG=350 LLG=355], and four molecules [RFZ=10.3 TFZ=10.7 PAK=0 LLG=164 RFZ=9.4 TFZ=22.0 PAK=4 LLG=510 RFZ=10.3 TFZ=29.2 PAK=7 LLG=1108 RFZ=8.5 TFZ=30.9 PAK=8 LLG=1727] in the asymmetric unit for clade C_{1086}, $E_{93TH057}$, and $C_{ZM109F.PB4}$, respectively. The initial map generated after rigid-body refinement was of sufficient quality to identify residues that differ from the search model. The map of the clade $E_{93TH057}$ was of such good quality that unmodeled electron density for a HEPES molecule, originated from the crystallization buffer, was clearly identifiable (Fig. S4). NCS averaging, solvent flipping, and density modification were applied in refinement, followed by group/individual ADP and TLS refinement using PHENIX. For refinement, NCS restraints, detected automatically by PHENIX, were employed for the initial refinement and released in the final stage of refinement for clade C_{1086}, but the restraints were maintained throughout entire refinement cycles for the lower resolution structures, clade B_{YU2} and clade C_{ZM109}, which provided us better stereochemistry and geometry of the models than releasing the restraints in refinement. Refinement statistics are summarized in Table S1.

Isothermal titration calorimetry. Isothermal titration calorimetry experiments were performed using an iTC_{200} Microcalorimeter from MicroCal Inc. (Northampton, MA). The calorimetric cell (~ 300 μl), containing wild-type gp120 or gp120 $core_e$ (44-492

ΔV12/V3) dialyzed in PBS (GIBCO) buffer, was titrated with sCD4, 17b Fab, or VRC01 Fab dialyzed in the same buffer by using 20 injections of 2 μl per injection. The concentration of gp120 full-length or $core_e$ was 3-5 μM and the concentration of the ligands in the injection syringe was between 50 and 75 μM. Data were analyzed with Microcal ORIGIN software using a single-site binding model. Protein concentrations were determined by absorbance at 280 nm based on extinction coefficients calculated from the protein sequences. All experiments were carried out at 25°C and 37°C.

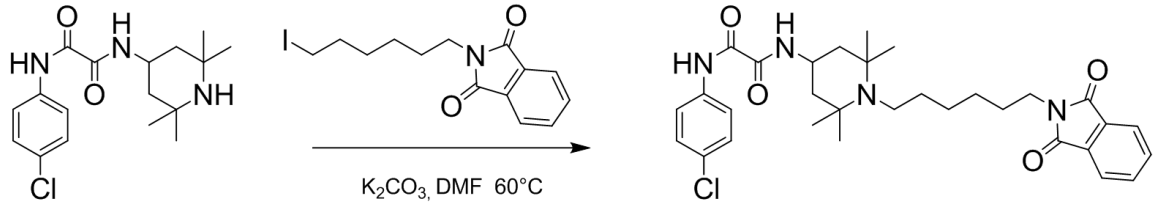
Expression and cross-linking of envelope glycoproteins. To generate radiolabeled HIV-1 gp120 proteins, 2×10^6 293T cells were transfected by the calcium phosphate method with the following codon-optimized vectors: pcDNA3.1-HIV-1 YU2, pcDNA3.1-HIV-1 YU2 44-492, pcDNA3.1-HIV-1 YU2 ΔV1/V2, pcDNA3.1-HIV-1YU2 ΔV3, pcDNA3.1-HIV-1 ΔV1/V2/V3 and pcDNA3.1-HIV-1 44-492ΔV1/V2/V3 ($core_e$), as previously described (6). Beginning one day after transfection, the cells were metabolically labeled for 16 h with 100 μCi/mL [35S]-methionine/cysteine ([35S] protein labeling mix (Perkin Elmer, Waltham, MA, USA) in Dulbecco's modified Eagle's medium lacking methionine and cysteine and supplemented with 5% dialyzed fetal bovine serum. The harvested culture supernatants were centrifuged at low speed to clear the cell debris, and the envelope glycoproteins were cross-linked as previously described (7). Briefly, clarified culture supernatants were incubated with 5 mM glutaraldehyde (GA) (Sigma) for 5 min at room temperature before addition of 50 mM glycine to quench the unreacted GA. The radiolabeled gp120 envelope glycoproteins (untreated or cross-linked) were precipitated with 2 μg of different anti-gp120 monoclonal antibodies in the presence of 50 μl of 10% Protein A-Sepharose (AmericanBioSciences, Blauvelt, NY, USA) for 1 hr at 37°C. All precipitated proteins were boiled for 5 min before being analyzed on NuPAGE Novex Bis-Tris polyacrylamide gels (Invitrogen, Carlsbad, CA, USA) under non-reducing conditions, as previously described (6).

Viral stocks and neutralization assays. HIV-1 Env-pseudoviruses were prepared by transfecting 293T cells (6 x 10^6 cells in 50 ml growth medium in a T-175 culture flask) with 10 μg of *rev*/*env* expression plasmid and 30 μg of an *env*-deficient HIV-1 backbone vector (pSG3Δ*env*), using Fugene 6 transfection reagent (Invitrogen). Pseudovirus-containing culture supernatants were harvested two days after transfection, filtered (0.45 μm), and stored at –80°C or in the vapor phase of liquid nitrogen. Neutralization was measured using HIV-1 Env-pseudoviruses to infect TZM-bl cells as described previously (8-11) with minor modifications. Briefly, the test reagent (sCD4, or test antibody, or NBD-556) was diluted in complete media containing 10% DMSO. Then 40 μl of virus was incubated for 30 min at 37ºC with 10 μl of serial diluted test reagent in duplicate wells of a 96-well flat bottom culture plate. To keep assay conditions constant, sham media containing 10% DMSO was used in place of test reagent in specified control wells. The virus input was set at a multiplicity of infection of approximately 0.01-0.1, which generally results in 100,000 to 400,000 relative light units (RLU) in a luciferase assay (Promega, Madison, WI). The test reagent concentrations were defined at the point of incubation with virus supernatant. The molecular masses of 46 kDa for sCD4 and 150 kDa for test antibodies were used to convert concentrations from mg/ml to μM. Neutralization curves were fit by nonlinear regression using a 5-parameter hill slope

equation as previously described (10). The 50% inhibitory concentrations (IC_{50}) were reported as the antibody concentrations required to inhibit infection by 50%.

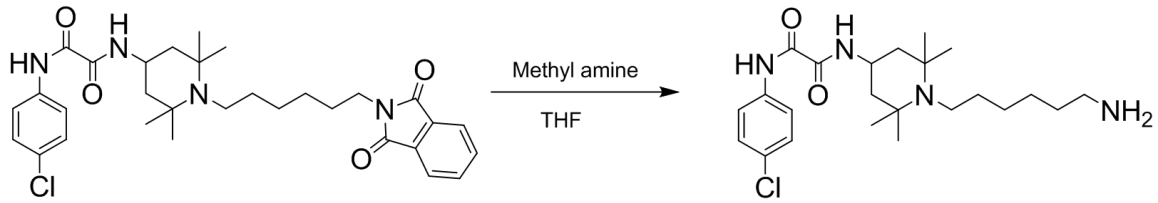
Surface plasmon resonance (SPR) experiments. All SPR experiments were performed using Biacore T100 optical biosensor (GE Healthcare) at 25ºC. Ligands (NBD556-hexylamine or d1d2 domain of sCD4) were immobilized (~6.5 or ~200 RU, respectively) on a CM5 sensor chip using standard amine coupling chemistry. The running buffer contained 10 mM Tris-HCl, pH 7.5, 350 mM NaCl, 0.02 % NaN_3, and 0.05% Surfactant P-20 (GE Healthcare). 10 mM glycine (pH 1.5) was used for regeneration of the surface of the CM5 sensor chip. For kinetic analysis, gp120 variants in an eight-sample concentration series using two-fold dilutions starting at 10 μM were injected at 40 μl/min. The association and dissociation phases were monitored for 60 seconds and 120 seconds, respectively. All kinetic data were extracted by fitting the responses globally with a 1:1 interaction model using Biacore T100 Evaluation software. For competition assay, ~4000 RU of anti-human IgG (Fc) antibody (GE Healthcare, Human Antibody Capture Kit) were immobilized on a CM5 chip. ~500 RU of soluble sCD4 were captured on the antibody and 100 nM of YU2 gp120 $core_e$ or gp120 full-length with 0, 5, 10, 25, 50, or 100 μM of NBD-556 were injected onto it at 40 μl/min. HBS-EP (GE Healthcare) buffer with 5% DMSO was used as the running buffer and all analytes for the competition assay were prepared to contain the same buffer. The association and dissociation phases were monitored for 120 seconds and 300 seconds, respectively, for the competition.

Synthesis of NBD556-hexylamine. All reactions were performed under a positive flow of nitrogen in flasks that were flame dried under vacuum. Aldrich SureSeal® anhydrous THF, DMF, and methylamine (33% in absolute ethanol) were used without further purification and stored under nitrogen. Silica gel (particle size 35-70 μm, 220-440 mesh) was used for column chromatography. Analytical samples were further purified by reverse-phase HPLC (Symmetry Shield™ C18, 150 x 19 mm, 5 μm, detection at 254 and 280 nm) eluting with a linear gradient of 20–90% MeOH in 0.05% TFA/H_2O. High-resolution electrospray ionization (ESI) mass data were measured on a Waters LCT Premier time-of-flight (TOF) mass spectrometer. NMR experiments were conducted on Bruker AVANCE 400-MHz spectrometer using specified solvents at 298 K. FT-IR data were acquired using a Perkin Elmer Spectrum One FTIR Spectrometer with a Universal ATR Sampling attachment. N-(ω-Iodohexyl) phthalimide (12) and NBD-556 (13) were prepared using previously published methods.



NBD-556 (0.15g, 0.4 mmol) and N-(ω-Iodohexyl) phthalimide (0.71 g, 2 mmol) were dissolved in 5ml of DMF followed by the addition of potassium carbonate (0.11 g, 0.8 mmol). The reaction mixture was left to stir at 60°C for 24 hrs. After cooling to room

temperature, water was added to the reaction. The mixture was then extracted with EtOAc and CH_2Cl_2.The organics were dried, concentrated and flash columned on silica gel with a gradient of 50% EtOAc/hexanes to 100% EtOAc giving the title compound as white solid in 35% yield. **^{1}H (400 MHz, CD_3OD)** δ 7.88-7.77 (m, 4H), 7.43 (d, *J*= 8.4 Hz, 2H), 7.28 (d, *J*= 8.4 Hz, 2H), 4.03 (m, 1H), 3.77 (t, *J*= 7.1 Hz, 2H), 3.64 (t, *J*= 7.4 Hz, 2H), 1.85-1.46 (m, 8H), 1.45-1.33 (m, 14H), 1.30 (t, J= 12.8Hz, 2H); **^{13}C (101 MHz, CD_3OD)** δ 170.0, 166.0, 165.0, 140.8, 135.5, 135.3, 133.5, 130.8, 130.7, 124.2, 58.5, 41.4, 41.3, 38.8, 30.7, 29.4, 28.3, 27.5, 27.2, 25.0; **IR (film):** ν_{max}= 3061, 2940, 1773, 1712, 1672, 1493, 1397, 1203, 1134, 1039; **HR-ESI-MS** 567.2733 *m/z* $[M+H]^+$ (calculated for $C_{23}H_{37}ClN_4O_2$, 566.2660).



The NBD-556-phthalimide (0.5 g, 0.9 mmol) was dissolved in 5 ml THF. 1.5 ml of methyl amine in ethanol was added then left to stir at RT overnight. Reaction mixture was then evaporated under reduced pressure and columned on silica gel initially in 30% MeOH/CH_2Cl_2 to remove impurities then ammonium hydroxide was added to the eluting solvent to make up 3% of the total volume. This solvent system elutes the title compounds as a colorless oil in 68% yield. **^{1}H (400 MHz, CD_3OD)** δ 7.45 (d, *J*= 8.6 Hz, 2H), 7.30 (d, *J*= 8.6 Hz, 2H), 4.03 (m, 1H), 3.79 (t, *J*= 7.2 Hz, 2H), 2.90 (t, *J*= 7.6 Hz, 2H), 1.71-1.42 (m, 8H), 1.45-1.34 (m, 14H), 1.28 (t, J= 13.0Hz, 2H); **^{13}C (101 MHz, CD_3OD)** δ 166.1, 165.1, 140.7, 135.5, 130.8, 58.5, 41.4, 41.3, 40.7, 30.7, 28.6, 28.2, 27.3, 27.1, 25.0; **IR (neat):** ν_{max}= 3064, 2945, 2866, 1669, 1645, 1493, 1202, 1135,; **HR-ESI-MS** 437.2685 *m/z* $[M+H]^+$ (calculated for $C_{23}H_{37}ClN_4O_2$, 436.2605).

Molecular morphs. Structural transitions between unliganded gp120 core and antibody-bound states were modeled using adiabatic mapping [14]. Coordinates were linearly interpolated over 100 frames with each subject to 30 minimization steps with CNS [15]. Frames were rendered using the PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC. and morphs displayed at 24 frames per second.

1. Zhou T, et al. (2007) Structural definition of a conserved neutralization epitope on HIV-1 gp120. *Nature* 445:732-737.

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

3. McCoy AJ, et al. (2007) Phaser crystallographic software. *J. Appl. Crystallogr*. 40:658-674.

4. Adams PD, et al. (2002) PHENIX: building new software for automated

crystallographic structure determination. *Acta Crystallogr. D* 58:1948-1954.

5. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* 60:2126-2132.

6. Finzi A, et al. (2010) Conformational characterization of aberrant disulfide-linked HIV-1 gp120 dimers secreted from overexpressing cells. *J. Virol. Methods* 168(1-2):155-161.

7. Yuan W, et al. (2006) Characterization of the multiple conformational states of free monomeric and trimeric human immunodeficiency virus envelope glycoproteins after fixation by cross-linker. *J. Virol.* 80:6725-6737.

8. Li M, et al. (2005) Human Immunodeficiency Virus Type 1 env Clones from Acute and Early Subtype B Infections for Standardized Assessments of Vaccine-Elicited Neutralizing Antibodies. *J. Virol*. 79:10108-10125.

9. Li M, et al. (2006) Genetic and neutralization properties of subtype C human immunodeficiency virus type 1 molecular env clones from acute and early heterosexually acquired infections in Southern Africa. *J. Virol*. 80:11776-11790.

10. Seaman MS, et al. (2010) Tiered categorization of a diverse panel of HIV-1 Env pseudoviruses for assessment of neutralizing antibodies. *J. Virol*. 84:1439-1452.

11. 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.

12. Lukyanenko NG, et al. (2004) Synthesis of lariat diazacrown ethers with terminal amino groups in the side chains. *Chem. Heterocycl. Compd.* 40:343-350.

13. Schon A, et al. (2006) Thermodynamics of binding of a low-molecular-weight CD4 mimetic to HIV-1 gp120. *Biochemistry* 45:10973-10980.

14. Krebs WG, Gerstein (2000) The morph server: a standardized system for analyzing and visualizing macromolecular motions in a database framework. Nucleic acids research 28:1665-1675.

15. Brunger AT, et al. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta crystallographica. Section D, Biological crystallography 54:905-921.

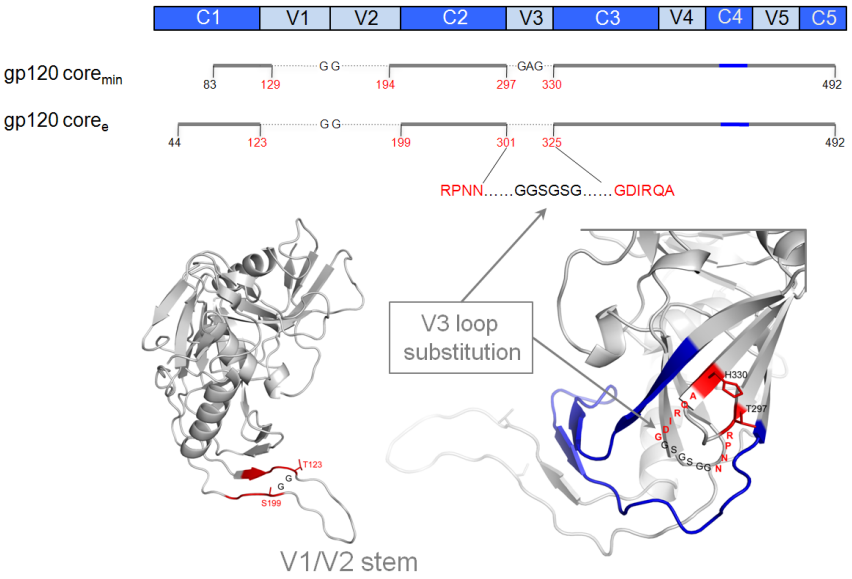


Fig. S1. Schematic diagram of the minimal core (core$_{min}$) and extended core (core$_e$) gp120 constructs. Core$_{min}$ spans residue 83 (N terminus) to 492 where the V1V2 region (129-194) is replaced with GG and the V3 region (297-330) is replaced with GAG. Core$_e$ spans residue 44 to 492. Its V1V2 region (123-199) is replaced with GG and the V3 region (301-325) is replaced with GGSGSG.

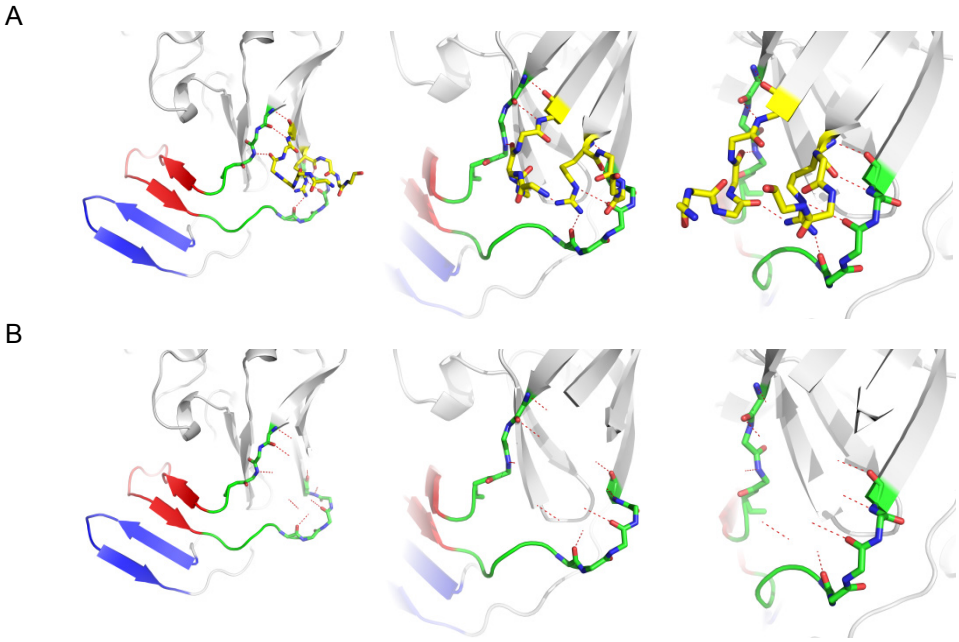


Fig. S2. Residues in gp120 core$_e$ V3 stem region (stick representation in yellow), which are missing in the core$_{min}$, form hydrogen bonds with the neighboring residues on the proximal and distal region of β20/21 (stick representation in green). (*A*) Close-up view of the V3 stem region of unliganded clade $E_{93TH057}$ core$_e$ from three different angles. Three pairs of hydrogen bonds in the proximal region and four pairs of hydrogen bonds in the distal region of β20/21 stabilize the V3 stem region. The structure and biochemical data suggest that absence of the residues (drawn in yellow) in the core$_{min}$, which is modeled in (*B*), destabilizes the region to effect a bridging sheet (drawn in red and blue) conformation different from that found in the unliganded core$_e$ and the CD4-bound gp120.

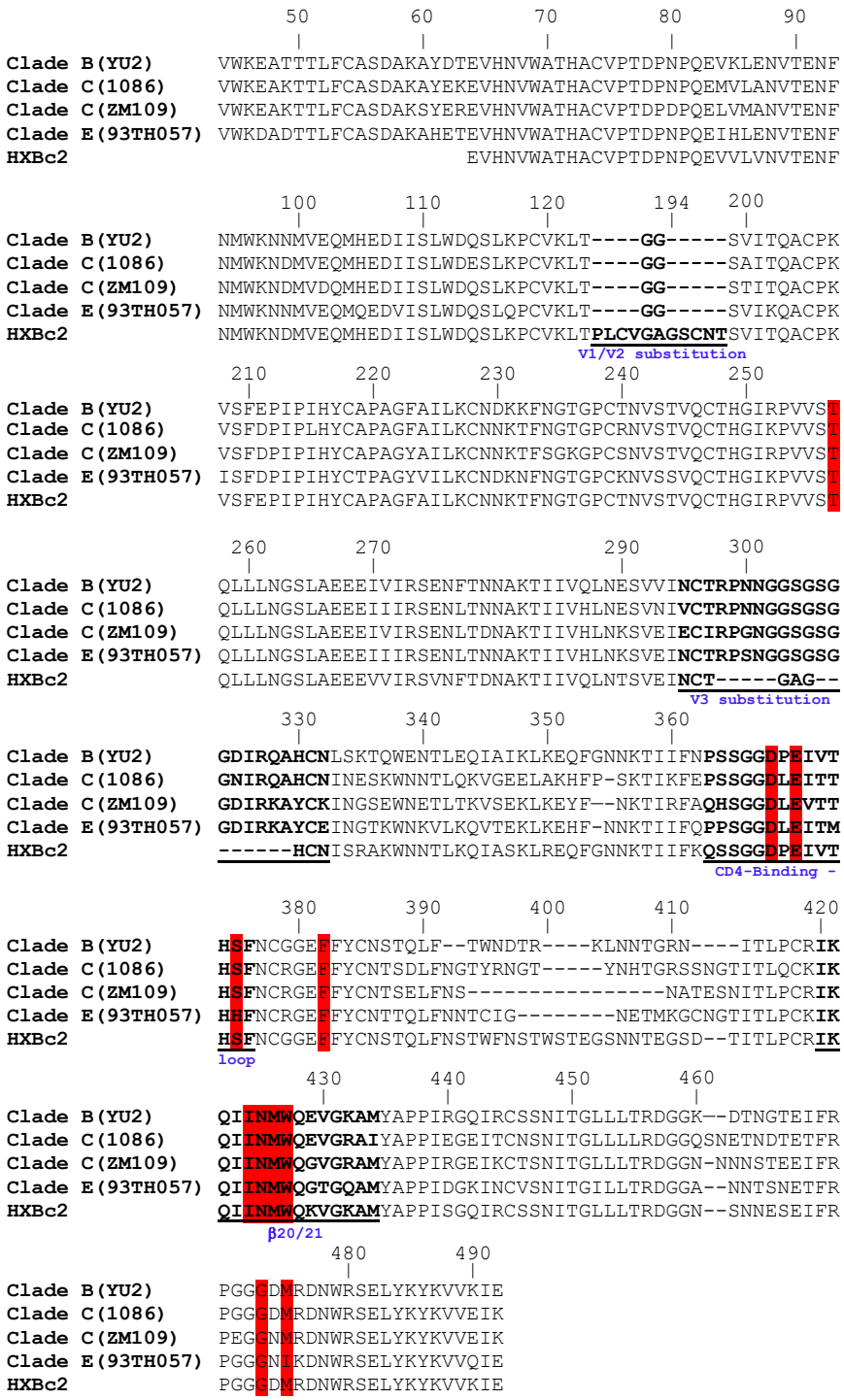


Fig. S3. Sequence alignment of HIV-1 clade B, C, and E gp120 cores. Clade B_{YU2}, clade C_{1086}, clade C_{ZM109}, and clade $E_{93TH057}$ gp120 $core_e$ sequences are aligned with HXBc2 gp120 $core_{min}$. Residues are numbered according to current convention, in reference to the HXBc2 numbers. The clade B_{YU2}, and clade C_{1086}, clade C_{ZM109}, and clade $E_{93TH057}$ $core_e$ constructs have a shorter V1/V2 stem and a longer V3 stem compared with those of the $core_{min}$ construct. The key residues lining the NBD-556 binding pocket are highlighted in red.

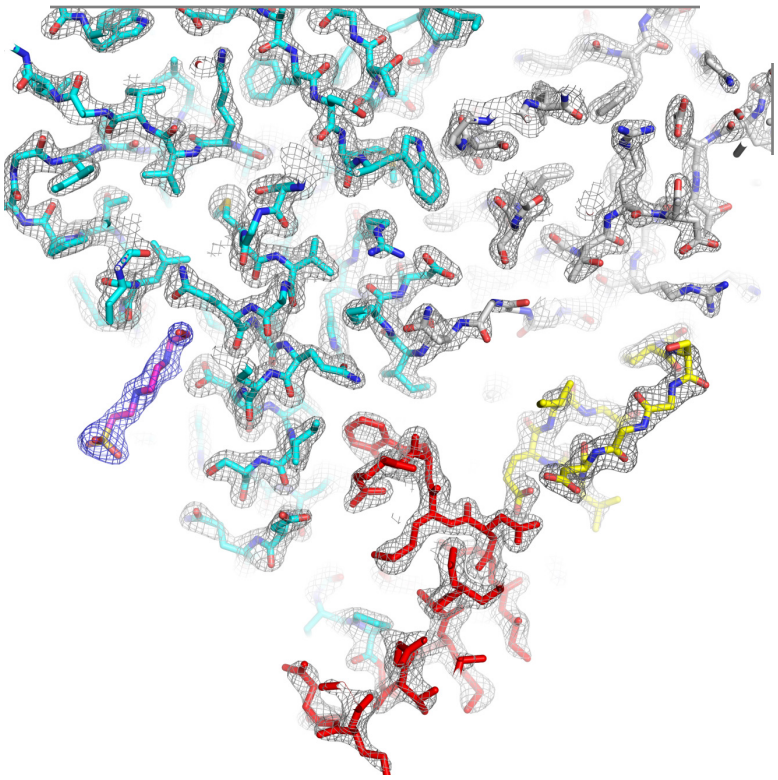


Fig. S4. *2fo-fc* map of clade $E_{93TH057}$ unliganded HIV-1 gp120 $core_e$ contoured at 1.5σ. Part of the inner domain, outer domain, CD4 binding loop, and bridging sheet of gp120 $core_e$ are colored in cyan, grey, yellow, and red, respectively. A HEPES molecule (purple) originated from the crystallization buffer occupies the well-defined electron density (blue mesh).

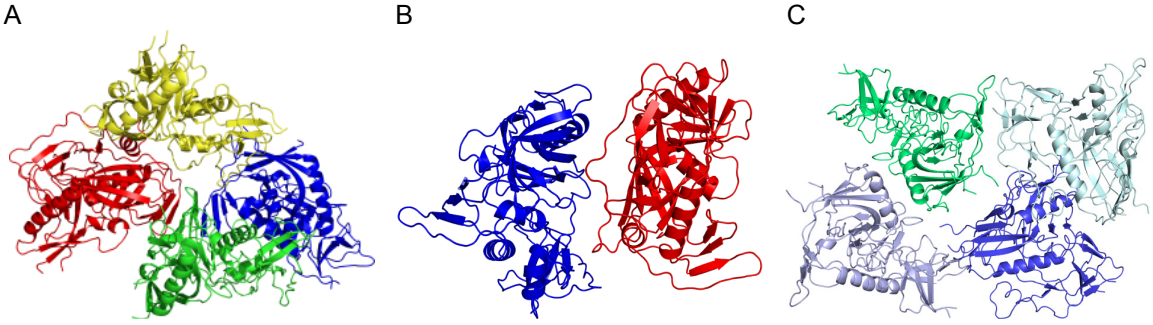


Fig. S5. Molecules found in the asymmetric unit of clade B_{YU2}, clade C_{1086}, and clade $C_{ZM109F.PB4}$ $core_e$ crystals. (*A*) A clade B_{YU2} $core_e$ crystal contains four gp120 molecules in the asymmetric unit in $P6_5$ packing. (*B*) The clade C_{1086} and (*C*) clade $C_{ZM109F.PB4}$ $core_e$ crystals contain two molecules in the asymmetric unit in $C222_1$ packing and four molecules in the asymmetric unit in $P2_1$ packing, respectively.

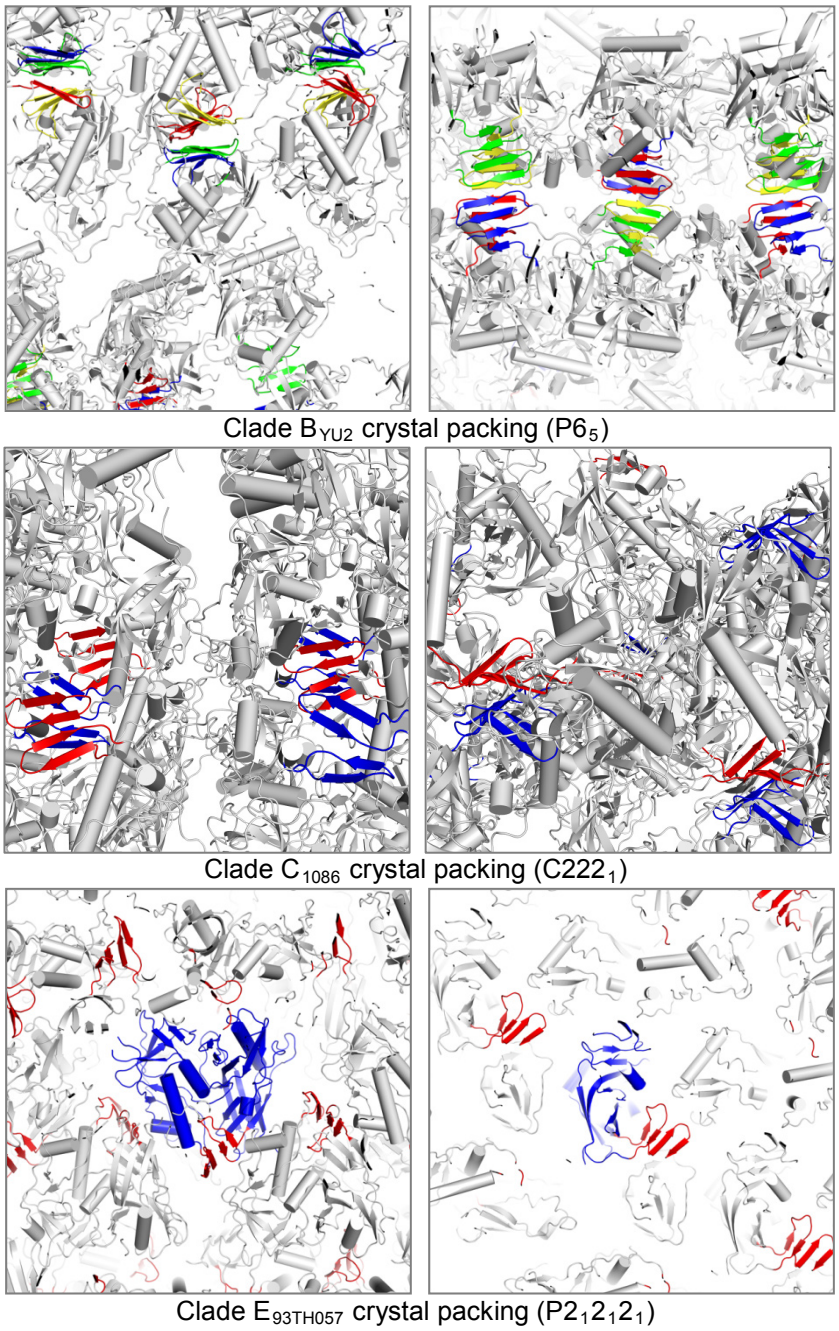


Fig. S6. Crystal packing of unliganded clade B, C, and E core$_e$. One of the surprising findings in the unliganded structures is the conformation of the bridging sheet, which is the same as that found in the CD4-bound conformation. Hence, we considered the possibility that the conformation was influenced by crystal lattice contacts. As shown in the $P6_5$ packing of clade B_{YU2} and $C222_1$ of clade C_{1086}, bridging sheets (highlighted in color) are in close proximity and making lattice contacts. The bridging sheets shown in $P2_12_12_1$ packing of the unliganded clade $E_{93TH057}$, however, do not make any lattice contacts. This suggests that the bridging sheet found in the unliganded structures enhanced the crystal lattice contacts, rather than the lattice contact inducing the conformation.

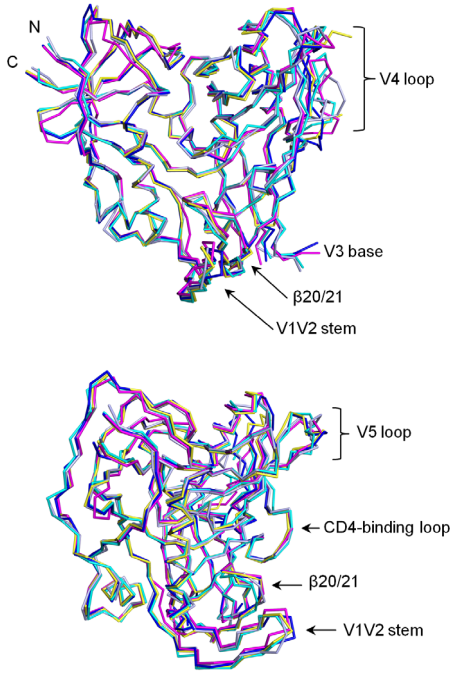


Fig. S7. Superposition of four unliganded and CD4-bound gp120 $core_e$ structures. The Cα trace of CD4- and 48d-bound (yellow), unliganded clade B_{YU2} (magenta), unliganded clade C_{1086} (blue), unliganded clade $C_{ZM109F.PB4}$ (light blue), and unliganded clade $E_{93TH057}$ (cyan) HIV-1 gp120 cores are superposed. Two perspectives are shown to highlight the most varied regions, the V4 and V5 loops.

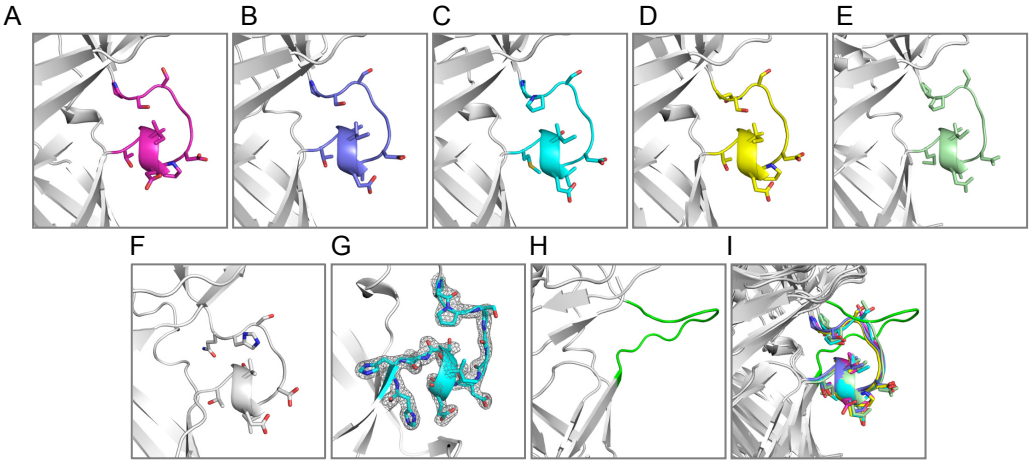


Fig. S8. CD4-binding loops of gp120s. (*A*) unliganded clade B_{YU2}, (*B*) unliganded clade C_{1086}, (*C*) unliganded clade $E_{93TH057}$, (*D*) CD4-and 48d-bound, (*E*) VRC01-bound, (*F*) unliganded clade $C_{ZM109F.PB4}$, (*G*) unliganded clade $E_{93TH057}$ in electron density contoured at 1.5σ, (*H*) unliganded SIV $core_{min}$, (*I*) superposition of the loops. All CD4-binding loops of HIV-1 unliganded gp120s are well superposed to those of CD4-bound and VRC01-bound gp120. Conformation of CD4-binding loop of unliganded SIV gp120, however, is markedly different.

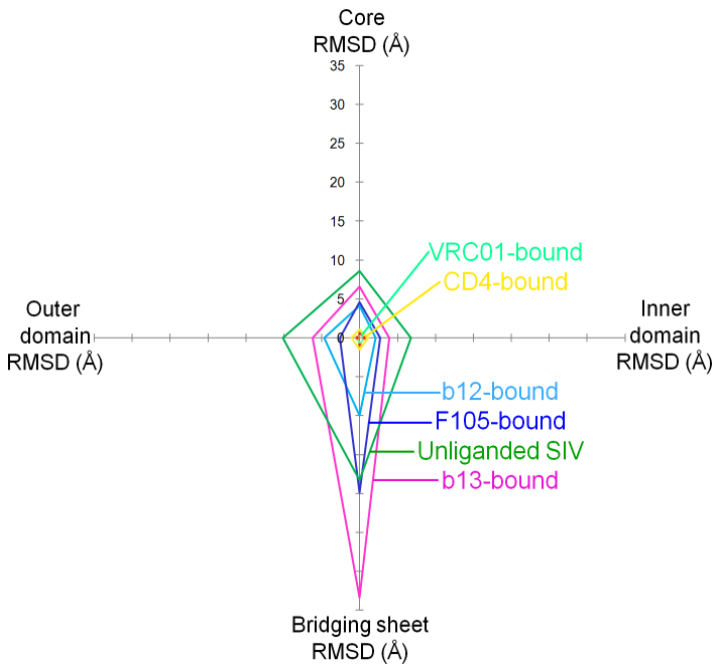


Fig. S9. Structural variation in HIV-1 gp120, relative to the unliganded clade E core$_e$ structure, overall and by domain. Each axis represents the coordinate deviation for the core, the inner domain, the outer domain, or the bridging sheet. Lines, drawn in the same color as the inner domains of the molecules shown in (Fig. 2*A*), display the rmsds for those molecules relative to the unliganded clade E core$_e$ structure.

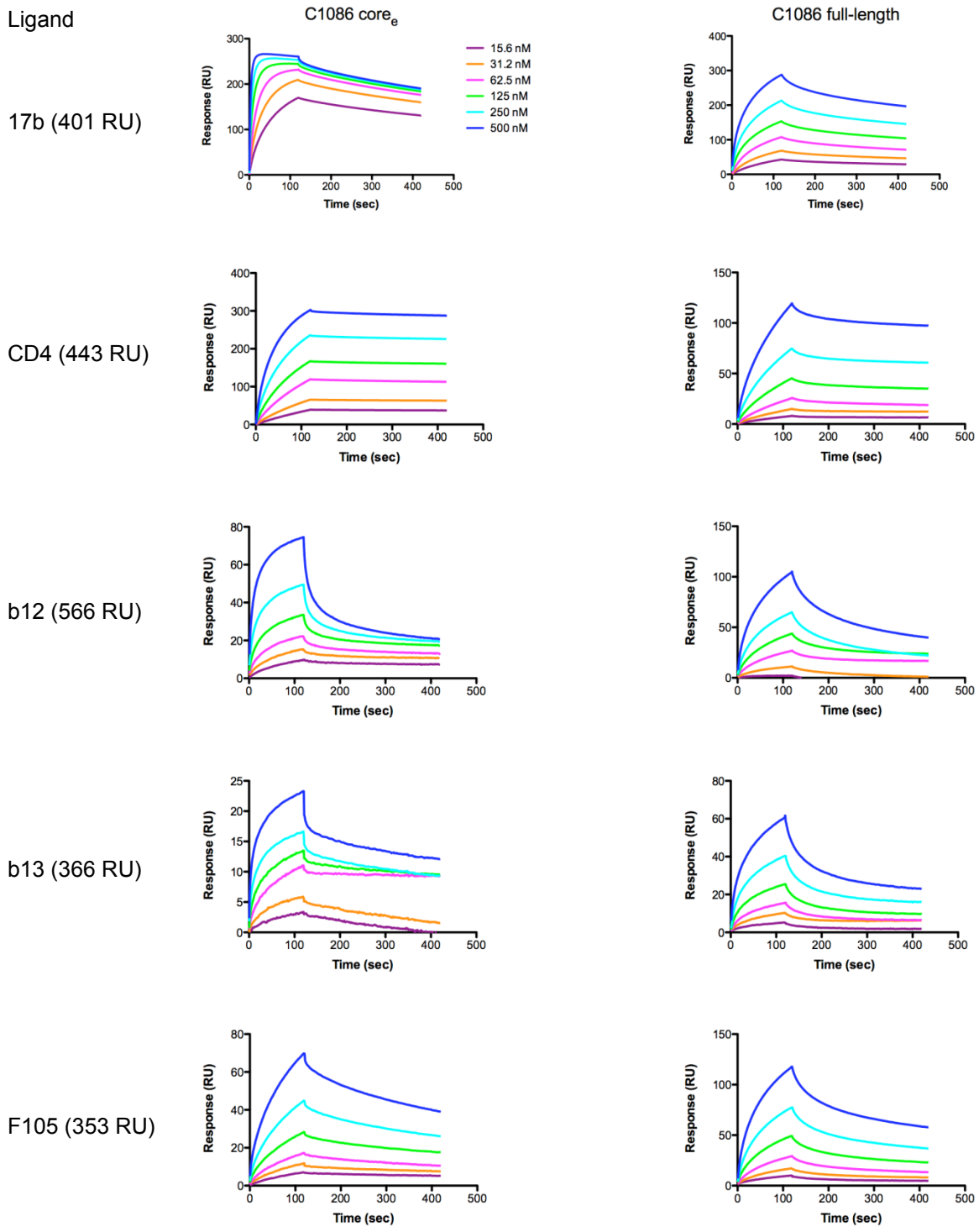


Fig. S10. SPR measurements of the 17b-purified C1086 core$_e$ and full-length gp120 binding to 17b and other antibodies. CD4 and antibodies were captured on anti-human IgG(Fc) antibody immobilized on a CM 5 chip, then core$_e$ and full-length gp120 in a six-concentration series using two-fold dilutions starting at 500 nM were passed over the sensor chip.

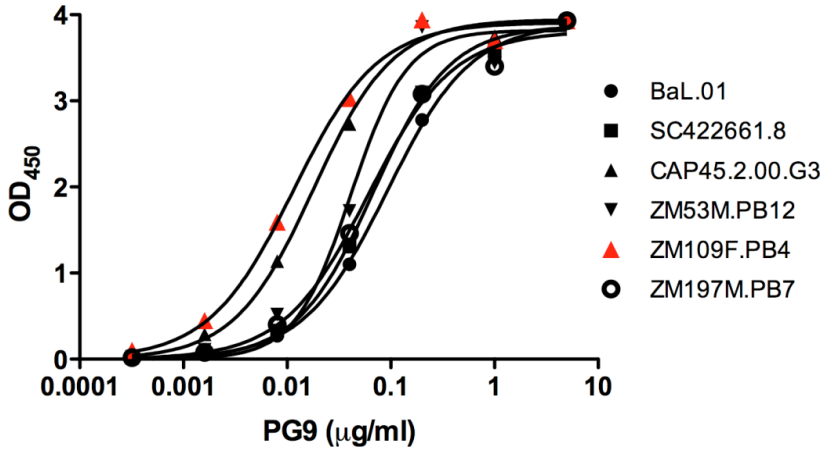


Fig. S11. High-affinity binding of PG9 to clade B and C gp120 monomers. Binding of PG9 to recombinant gp120 monomers from 30 clade B and C isolates was assessed by ELISA, and the binding of PG9 to the six monomers that showed the highest level of binding is shown here.

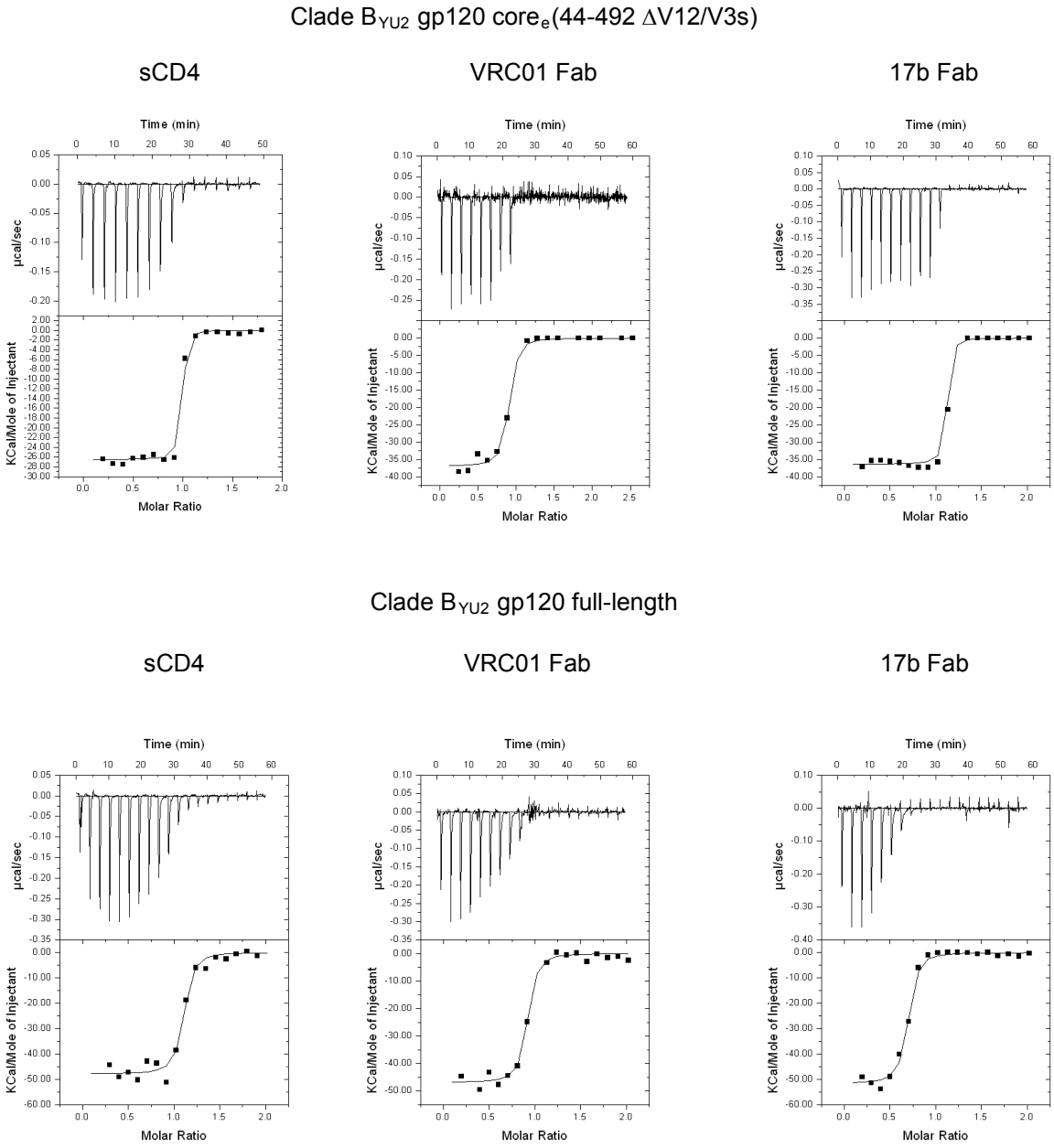


Fig. S12. Clade B_{YU2} gp120-ligand binding thermodynamics. Calorimetry data for the titration of full-length and $core_e$ gp120 variants with sCD4, VRC01 Fab, and 17b Fab in 10 mM Na_2HPO_4, 200 mM NaCl (pH 7.4). The top panels show raw data in power versus time. The area under each spike is proportional to the heat produced at each injection. The lower panels show integrated areas normalized to the number of moles of sCD4, 17b Fab, or VRC01 Fab injected at each injection step.

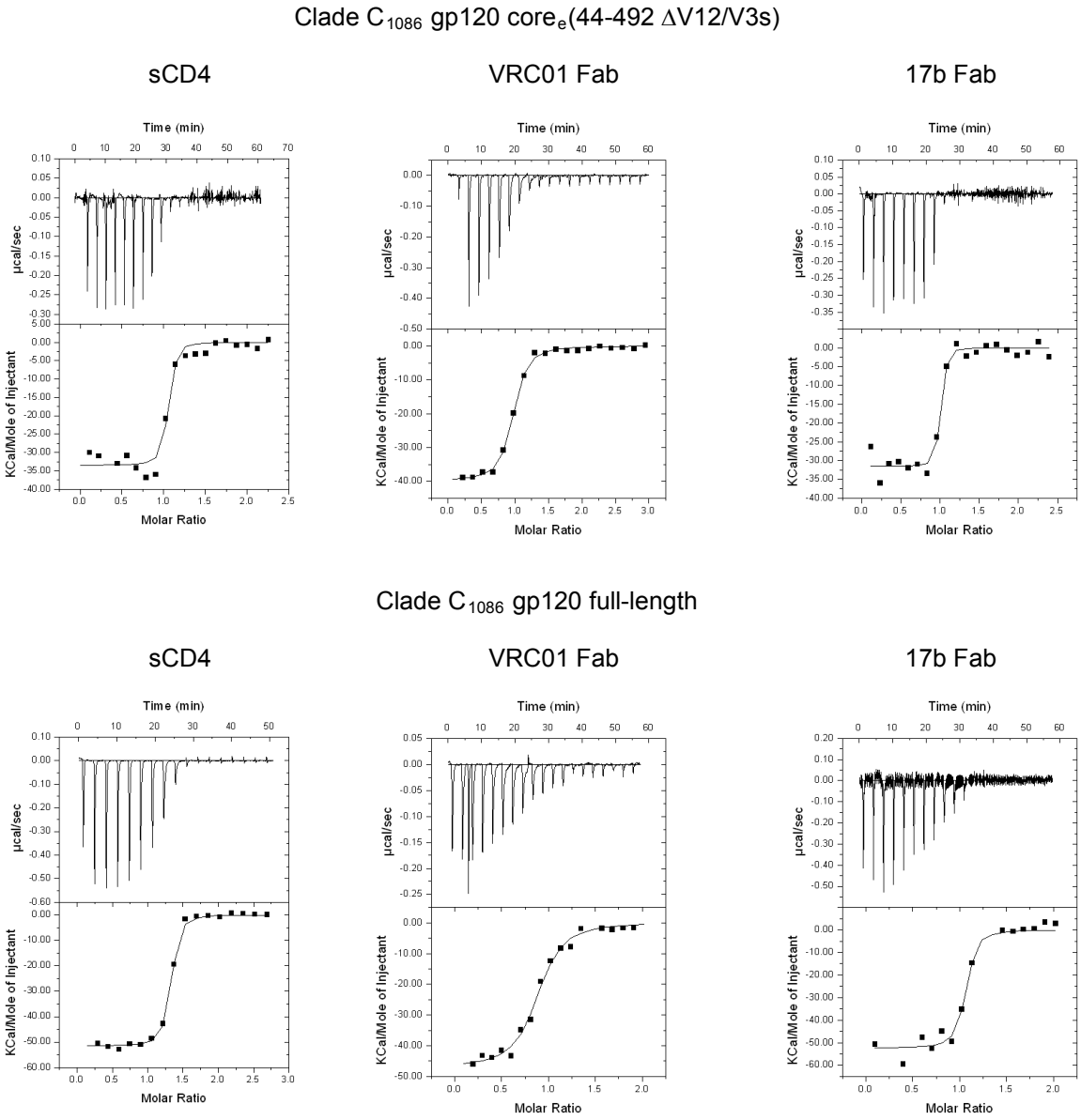


Fig. S13. Clade C_{1086} gp120-ligand binding thermodynamics. Calorimetry data for the titration of full-length and $core_e$ gp120 variants with sCD4, VRC01 Fab, and 17b Fab in 10 mM Na_2HPO_4, 200 mM NaCl (pH 7.4). The top panels show raw data in power versus time. The area under each spike is proportional to the heat produced at each injection. The lower panels show integrated areas normalized to the number of moles of sCD4, 17b Fab, or VRC01 Fab injected at each injection step.

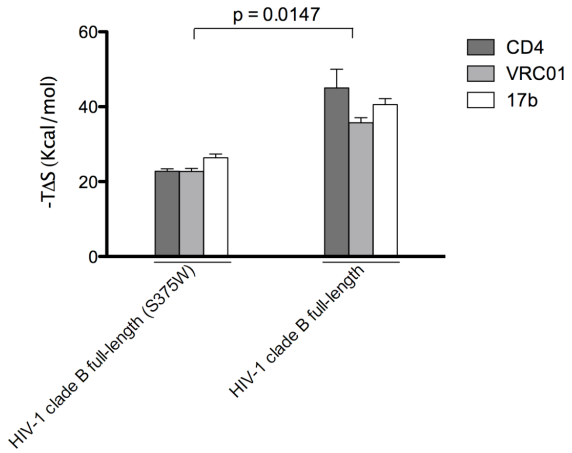


Fig. S14. Entropy of ligand interactions with HIV-1 clade B full-length gp120 and clade B full-length S375W. The entropy change on the S375W mutant gp120 interactions with CD4, VRC01, or 17b is about half of that observed for the full-length, wild-type gp120.

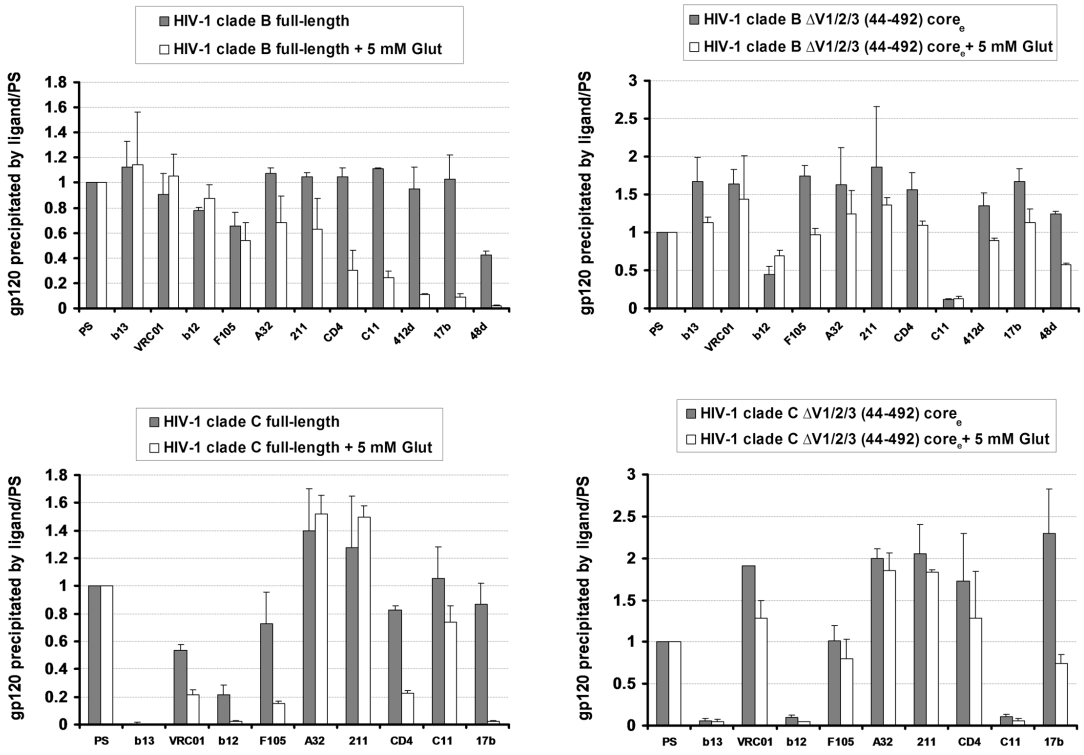


Fig. S15. Effect of loop deletion on antigen recognition. Radiolabeled clade B_{YU2} and clade C_{1086} full-length or $core_e$ HIV-1 gp120 glycoproteins were either untreated or cross-linked with 5 mM glutaraldehyde (Glut) before being precipitated with a panel of anti-gp120 monoclonal antibodies. The amount of protein precipitated was normalized to that precipitated by a polyclonal mixture of sera (PS) from HIV-1 infected individuals.

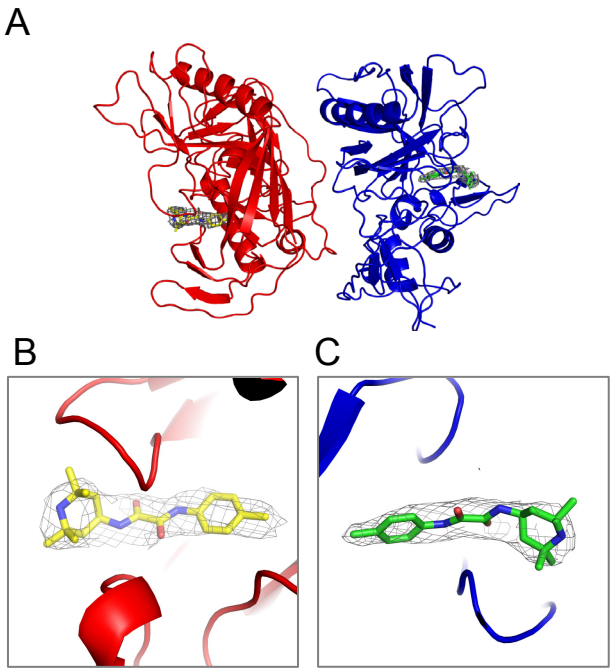


Fig. S16. Structure of NBD-556 bound C_{1086} gp120 $core_e$ in the asymmetric unit. (*A*) Two molecules of gp120 are found in the asymmetric unit of the C_{1086} gp120-NBD-556 crystals; these are colored red and blue and are shown in ribbon representation. Each of the "Phe43" binding cavities of the gp120 is filled with an NBD-556 molecule (displayed in stick representation). (*B*, *C*) Close-up of the NBD-556 molecules for each of the two copies in the asymmetric unit are shown with *2Fo-Fc* electron density contoured at 1σ.

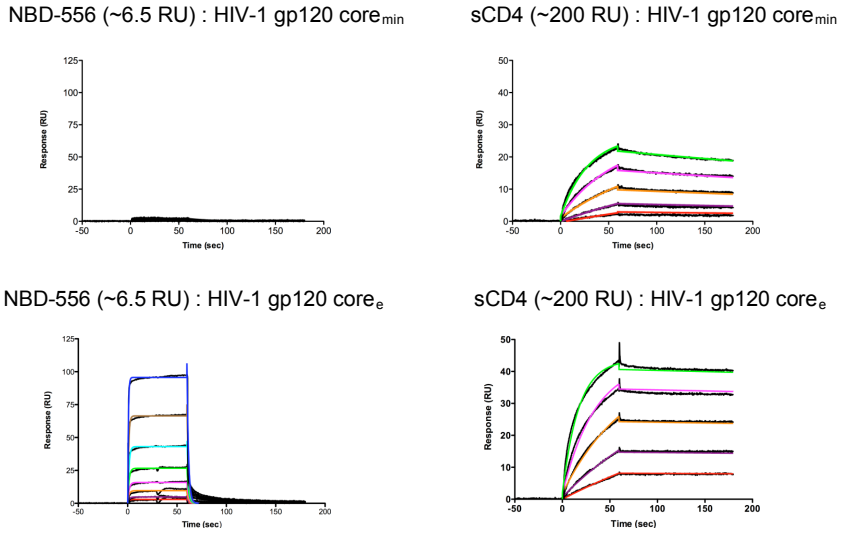


Fig. S17. Comparison of HIV-1 clade B_{YU2} gp120 $core_{min}$ interactions with NBD-556 or sCD4 and gp120 $core_e$ with NBD-556 or sCD4. A stark contrast exists between gp120 $core_{min}$ and $core_e$ for their binding responses to NBD-556, whereas both cores recognize sCD4 with similar K_D's. Colored lines represent the fit of the kinetic data to a 1:1 binding model. gp120 $core_{min}$ and $core_e$ were injected in an eight-sample concentration series using two-fold dilutions starting at 10 μM, and in a five-sample concentration series starting at 1.25 μM for sCD4.

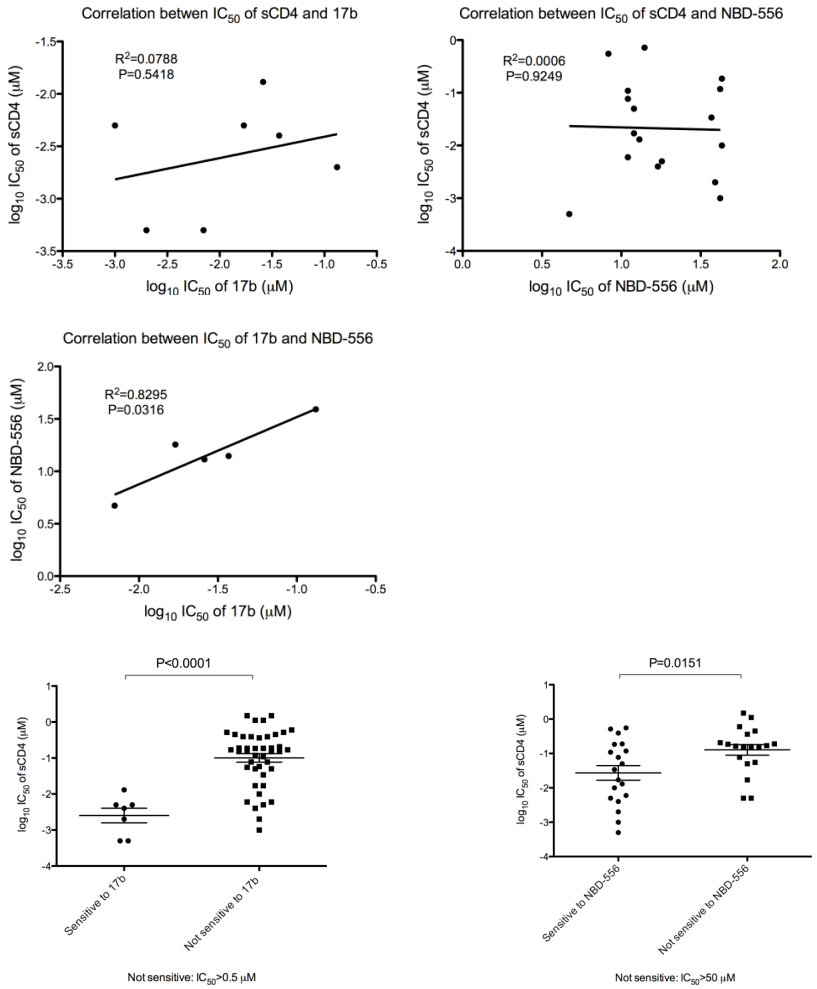


Fig. S18. Correlations between IC_{50} of sCD4, 17b, and NBD-556 against clade B and C viruses. The IC_{50} values of sCD4, NBD-556, and 17b against clade B and C reference viruses showed significant correlations only for NBD-556 and 17b. Viruses sensitive to 17b or NBD-556 were more sensitive to sCD4 than those not sensitive to 17b or NBD-556.

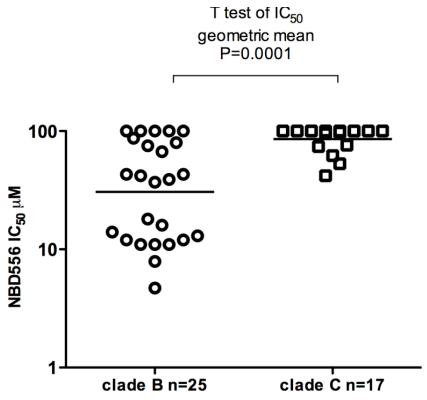


Fig. S19. Geometric mean of IC_{50} of NBD-556 against clade B and C reference viruses. Overall, clade B viruses are more sensitive to NBD-556 than clade C viruses.

Table S1. Crystallization screening results for gp120 core$_e$ proteins and co-crystallization with NBD-556 are summarized.

HIV-1 Strain	Expression	Purification	Crystals	Data (Å)
Clade B $_{YU2}$	293fs	17b/ConA/S200	X	3.4
Clade E$_{93TH057}$	293fs	17b/ConA/S200	X	1.9
AB253428[1]	293fs	17b/ConA/S200	X	--
C97ZA	293fs	17b/ConA/S200	X	--
VRC-A	293fs	17b/ConA/S200	--	--
AF082394[1]	293fs	17b/ConA/S200	--	--
EF165541[1]	GnTi$^-$	17b/ConA/S200	--	--
Clade C$_{1086}$	GnTi$^-$	17b/ConA/S200	X	2.8
Clade C$_{ZM109F.PB4}$	GnTi$^-$	17b/ConA/S200	X	4.0
AJ249239[1]	GnTi$^-$	17b	--	--
VRC-B	GnTi$^-$	17b/ConA/S200	--	--
AY046058[1]	GnTi$^-$	17b	--	--
AF457052[1]	GnTi$^-$	17b/ConA/S200	--	--
Clade C$_{1086}$:NBD556	GnTi$^-$	17b/ConA/S200	X	3.0

[1] Accession number 17b: 17b IgG conjugated-protein A column, ConA: concanavalin A column, S200: Superdex 200 size exclusion column.

Table S2. Summary of X-ray crystallographic data and model refinement statistics

Data Collection

gp120	Clade B_{YU2}	Clade C_{1086}	Clade $C_{ZM109F.PB4}$	Clade C_{1086}: NBD-556	Clade $E_{93TH057}$
Resolution (Å)	50-3.4(3.5-3.4)	50-2.8(2.9-2.8)	50-4.0(4.14-4.0)	45.3-2.7(2.8-2.7)	50-1.9(1.93-1.9)
Space group	$P6_5$	$C222_1$	$P12_11$	$C222_1$	$P2_12_12_1$
Wavelength (Å)	1.0	1.0	1.0	1.0	1.0
Molecule per ASU	4	2	4	2	1
Unit cell constants (Å)					
a	222.9	67.8	55.6	66.9	63.5
b	222.9	127.5	197.2	126.0	66.9
c	86.4	193.0	83.4	191.4	88.0
α, β, γ (°)	90, 90, 120	90, 90, 90	90, 91.3, 90	90, 90, 90	90, 90, 90
Completeness (%)	93.6(70.0)	93.8(82.4)	84.8(57.8)	74.0(26.3)	98.3(86.5)
No. of total reflections	186,054	84,043	42,517	91,321	208,725
No. of unique reflections	31,621	20,001	12,925	17,560	30,907
Redundancy	4.9(2.5)	4.2(2.5)	3.3(2.5)	5.2(3.6)	6.8(5.5)
R_{sym} (%)[1]	12.0	7.9	15.1	11.6	6.2
I/σ	15.9(2.0)	29.7(3.8)	6.9(2.1)	26.2(4.8)	38.5(2.0)

Refinement Statistics ($|F| > 0\ \sigma$)

Resolution (Å)	48.2-3.4	20-2.8	42-4.0	45.3-2.7	28.7-1.9
No. of reflections	31,870	19,503	11,816	17,156	29,250
$R_{cryst,}$(%)[2]	27.7	22.3	28.3	23.9	18.9
R_{free} (%)[3]	31.2	28.1	31.2	29.5	23.0
Average *B*-value	97.3	76.6	139.5	70.7	48.6
RMSD bonds (Å)	0.002	0.002	0.003	0.007	0.007
RMSD angles (°)	0.578	0.607	0.601	0.669	1.148
Ramachandran plot					
Favored (%)	95.1	95.4	92.2	91.7	96.7
Allowed (%)	99.7	99.9	99.3	100.0	100.0
PDB ID	3TGQ	3TGR	3TIH	3TGS	3TGT

[1] $R_{sym} = \Sigma|I - \langle I\rangle|/\Sigma\langle I\rangle$; I is the intensity of an individual measurement and <I> its mean value. [2] $R_{cryst} = \Sigma|F_o - F_c|/|F_o|$; F_o and F_c are observed and calculated structure factors, respectively. [3] R_{free}: same as R_{cryst} but calculated for a subset of the reflections (5%), which were omitted during the refinement and used to monitor its convergence.

Table S3. Structural comparison of unliganded clade B_{YU2} gp120 and gp120s in other crystalline lattices and bound by other ligands [1]

Clade	Ligand(s)	PDB	Chain	$RMSD_{Com}$ (Å)	$RMSD_{ID}$ (Å)	$RMSD_{OD}$ (Å)	$RMSD_{BS}$ (Å)
B (HXBc2)	CD4, 48d	3JWO	A	0.75	0.90	0.92	1.07
B (HXBc2)	CD4, 48d	3JWD	A	0.76	0.91	0.94	1.06
B (HXBc2)	CD4, 17b	2NXY	A	0.77	0.87	0.93	0.96
B (HXBc2)	CD4, 17b	2NXZ	A	0.76	0.88	0.94	0.91
B (YU2)	CD4, 412d	2QAD	A	0.73	0.93	0.72	1.32
B (HXBc2)	CD4, 17b	1G9M	G	0.83	1.09	0.89	0.90
B (HXBc2)	CD4, 17b	1G9N	G	0.72	0.80	0.92	0.62
B (JR-FL)	CD4, X5	2B4C	G	0.79	0.95	0.93	0.88
B (YU2)	F23, 17b	1YYM	G	0.78	0.86	1.04	1.04
B (HXBc2)	CD4, 17b	1RZJ	G	0.82	1.07	0.97	0.90
B (YU2)	CD4, 17b	1RZK	G	0.74	0.81	0.94	0.63
B (YU2)	CD4M33, 17b	1YYL	G	0.82	0.92	1.06	1.09
B (YU2)	[Phe23]M47, 17b	2I60	G	0.76	0.88	1.05	0.98
B (YU2)	CD4M47, 17b	2I5Y	G	0.81	0.94	1.05	0.95
B (HXBc2)	CD4, 17b	1GC1	G	1.32	1.51	1.59	1.73
C (CAP210)	CD4, 21c	3LQA[2]	G	2.03	1.34	2.48	1.42
B (HXBc2)	b12	2NY7	G	4.25	4.85	2.16	10.33
B (YU2)	F105	3HI1	G	4.78	2.71	2.77	20.34
B (HXBc2)	b13	3IDX	G	6.60	6.26	3.86	33.76
SIV	Unliganded	2BF1[2]	A	8.42	9.89	6.89	17.44
C1086	Unliganded	3TGR	A	0.80	0.97	0.72	1.14
E(93TH057)	Unliganded	3TGT	A	1.12	1.10	1.16	1.60
C(ZM109)	Unliganded	3TIH	A	1.16	1.05	1.25	1.46
E(93TH057)	VRC01	3NGB	A	1.12	1.09	1.08	1.91

[1] Cα-RMSDs of a set of common residues shared by all gp120s ($RMSD_{Com}$), of inner domain ($RMSD_{ID}$), of outer domain ($RMSD_{OD}$) and of bridging sheet ($RMSD_{BS}$) were calculated after each of the gp120 structures was superimposed on the VRC01-bound gp120 structure. The common set consists of 253 residues, including residues 90-120, 204-205, 215-298, 329-355, 357-392, 395, 413-469 and 473-489. The inner domain (ID) contains three segments, residues 83-118, 206-252 and 476-492. The outer domain (OD) contains three segments, residues 253-397, 410-419 and 437-475. The bridging sheet contains two regions, residues 118-205 and 421-437.

[2] For 3LQA, which has adopted a different residue numbering, and 2BF1, which is the gp120 from SIV, sequence alignment with 1G9M is used to derive the standard HXBc2 numbering.

Table S4. Structural comparison of unliganded clade C_{1086} gp120 and gp120s in other crystalline lattices and bound by other ligands [1]

Clade	Ligand(s)	PDB	Chain	$RMSD_{Com}$ (Å)	$RMSD_{ID}$ (Å)	$RMSD_{OD}$ (Å)	$RMSD_{BS}$ (Å)
B (HXBc2)	CD4, 48d	3JWO	A	0.57	0.53	0.68	0.83
B (HXBc2)	CD4, 48d	3JWD	A	0.58	0.54	0.69	0.82
B (HXBc2)	CD4, 17b	2NXY	A	0.70	0.94	0.67	1.07
B (HXBc2)	CD4, 17b	2NXZ	A	0.71	0.94	0.67	1.06
B (YU2)	CD4, 412d	2QAD	A	0.84	1.09	0.66	1.11
B (HXBc2)	CD4, 17b	1G9M	G	0.79	1.16	0.73	1.13
B (HXBc2)	CD4, 17b	1G9N	G	0.76	0.96	0.78	1.06
B (JR-FL)	CD4, X5	2B4C	G	0.68	0.90	0.66	0.72
B (YU2)	F23, 17b	1YYM	G	0.83	0.99	0.86	0.92
B (HXBc2)	CD4, 17b	1RZJ	G	0.76	1.14	0.71	1.11
B (YU2)	CD4, 17b	1RZK	G	0.78	0.97	0.80	1.05
B (YU2)	CD4M33, 17b	1YYL	G	0.88	1.03	0.91	1.03
B (YU2)	[Phe23]M47, 17b	2I60	G	0.81	0.96	0.87	0.91
B (YU2)	CD4M47, 17b	2I5Y	G	0.87	1.01	0.93	0.95
B (HXBc2)	CD4, 17b	1GC1	G	1.34	1.55	1.25	1.99
C (CAP210)	CD4, 21c	3LQA[2]	G	1.96	1.17	2.42	1.53
B (HXBc2)	b12	2NY7	G	4.14	4.75	1.92	10.05
B (YU2)	F105	3HI1	G	4.68	2.75	2.56	20.24
B (HXBc2)	b13	3IDX	G	6.47	6.10	3.57	33.54
SIV	Unliganded	2BF1[2]	A	8.44	9.94	6.79	18.11
B (YU2)	Unliganded	3TGQ	A	0.80	0.97	0.72	1.14
E(93TH057)	Unliganded	3TGT	A	0.92	0.61	0.95	1.07
C(ZM109)	Unliganded	3TIH	A	1.03	0.77	1.17	1.03
E(93TH057)	VRC01	3NGB	A	0.90	0.63	0.86	1.17

[1] Cα-RMSDs of a set of common residues shared by all gp120s ($RMSD_{Com}$), of inner domain ($RMSD_{ID}$), of outer domain ($RMSD_{OD}$) and of bridging sheet ($RMSD_{BS}$) were calculated after each of the gp120 structures was superimposed on the VRC01-bound gp120 structure. The common set consists of 253 residues, including residues 90-120, 204-205, 215-298, 329-355, 357-392, 395, 413-469 and 473-489. The inner domain (ID) contains three segments, residues 83-118, 206-252 and 476-492. The outer domain (OD) contains three segments, residues 253-397, 410-419 and 437-475. The bridging sheet contains two regions, residues 118-205 and 421-437.

[2] For 3LQA, which has adopted a different residue numbering, and 2BF1, which is the gp120 from SIV, sequence alignment with 1G9M is used to derive the standard HXBc2 numbering.

Table S5. Structural comparison of unliganded clade $E_{93TH057}$ gp120 and gp120s in other crystalline lattices and bound by other ligands [1]

Clade	Ligand(s)	PDB	Chain	$RMSD_{Com}$ (Å)	$RMSD_{ID}$ (Å)	$RMSD_{OD}$ (Å)	$RMSD_{BS}$ (Å)
B (HXBc2)	CD4, 48d	3JWO	A	0.91	0.55	1.19	0.82
B (HXBc2)	CD4, 48d	3JWD	A	0.91	0.55	1.20	0.83
B (HXBc2)	CD4, 17b	2NXY	A	1.03	0.99	1.10	1.21
B (HXBc2)	CD4, 17b	2NXZ	A	1.06	1.00	1.11	1.30
B (YU2)	CD4, 412d	2QAD	A	1.05	1.18	1.09	0.95
B (HXBc2)	CD4, 17b	1G9M	G	1.11	1.25	1.11	1.32
B (HXBc2)	CD4, 17b	1G9N	G	1.08	1.11	1.12	1.45
B (JR-FL)	CD4, X5	2B4C	G	1.10	1.10	1.22	1.04
B (YU2)	F23, 17b	1YYM	G	1.08	1.06	1.18	1.20
B (HXBc2)	CD4, 17b	1RZJ	G	1.10	1.24	1.16	1.31
B (YU2)	CD4, 17b	1RZK	G	1.09	1.12	1.14	1.42
B (YU2)	CD4M33, 17b	1YYL	G	1.13	1.13	1.18	1.25
B (YU2)	[Phe23]M47, 17b	2I60	G	1.11	1.09	1.20	1.30
B (YU2)	CD4M47, 17b	2I5Y	G	1.15	1.16	1.28	1.34
B (HXBc2)	CD4, 17b	1GC1	G	1.53	1.55	1.61	2.18
C (CAP210)	CD4, 21c	3LQA[2]	G	2.16	1.27	2.62	1.65
B (HXBc2)	b12	2NY7	G	4.15	4.68	2.11	9.95
B (YU2)	F105	3HI1	G	4.59	2.60	2.67	19.91
B (HXBc2)	b13	3IDX	G	6.58	6.18	3.90	33.36
SIV	Unliganded	2BF1[2]	A	8.55	10.11	6.79	18.26
B (YU2)	Unliganded	3TGQ	A	1.12	1.10	1.16	1.60
C1086	Unliganded	3TGR	A	0.92	0.61	0.95	1.07
C(ZM109)	Unliganded	3TIH	A	0.81	0.67	0.98	0.89
E(93TH057)	VRC01	3NGB	A	0.51	0.32	0.60	0.88

[1] Cα-RMSDs of a set of common residues shared by all gp120s ($RMSD_{Com}$), of inner domain ($RMSD_{ID}$), of outer domain ($RMSD_{OD}$) and of bridging sheet ($RMSD_{BS}$) were calculated after each of the gp120 structures was superimposed on the VRC01-bound gp120 structure. The common set consists of 253 residues, including residues 90-120, 204-205, 215-298, 329-355, 357-392, 395, 413-469 and 473-489. The inner domain (ID) contains three segments, residues 83-118, 206-252 and 476-492. The outer domain (OD) contains three segments, residues 253-397, 410-419 and 437-475. The bridging sheet contains two regions, residues 118-205 and 421-437.

[2] For 3LQA, which has adopted a different residue numbering, and 2BF1, which is the gp120 from SIV, sequence alignment with 1G9M is used to derive the standard HXBc2 numbering.

Table S6. Structural comparison of unliganded clade $C_{ZM109F.PB4}$ gp120 and gp120s in other crystalline lattices and bound by other ligands [1]

Clade	Ligand(s)	PDB	Chain	$RMSD_{Com}$ (Å)	$RMSD_{ID}$ (Å)	$RMSD_{OD}$ (Å)	$RMSD_{BS}$ (Å)
B (HXBc2)	CD4, 48d	3JWO	A	0.99	0.63	1.40	0.90
B (HXBc2)	CD4, 48d	3JWD	A	1.00	0.63	1.42	0.89
B (HXBc2)	CD4, 17b	2NXY	A	1.12	1.05	1.34	1.05
B (HXBc2)	CD4, 17b	2NXZ	A	1.15	1.07	1.35	1.10
B (YU2)	CD4, 412d	2QAD	A	1.15	1.29	1.28	0.97
B (HXBc2)	CD4, 17b	1G9M	G	1.18	1.31	1.34	1.16
B (HXBc2)	CD4, 17b	1G9N	G	1.19	1.14	1.40	1.32
B (JR-FL)	CD4, X5	2B4C	G	1.11	1.08	1.30	1.02
B (YU2)	F23, 17b	1YYM	G	1.18	1.06	1.42	1.17
B (HXBc2)	CD4, 17b	1RZJ	G	1.17	1.30	1.37	1.16
B (YU2)	CD4, 17b	1RZK	G	1.20	1.15	1.42	1.31
B (YU2)	CD4M33, 17b	1YYL	G	1.21	1.13	1.41	1.24
B (YU2)	[Phe23]M47, 17b	2I60	G	1.20	1.09	1.44	1.22
B (YU2)	CD4M47, 17b	2I5Y	G	1.24	1.17	1.50	1.33
B (HXBc2)	CD4, 17b	1GC1	G	1.57	1.65	1.68	2.08
C (CAP210)	CD4, 21c	3LQA[2]	G	2.18	1.42	2.66	1.47
B (HXBc2)	b12	2NY7	G	4.23	4.89	2.31	9.96
B (YU2)	F105	3HI1	G	4.63	2.74	2.88	19.71
B (HXBc2)	b13	3IDX	G	6.58	6.22	3.98	33.40
SIV	Unliganded	2BF1[2]	A	8.51	10.09	6.77	17.90
C1086	Unliganded	3TGR	A	1.03	0.77	1.17	1.03
E(93TH057)	Unliganded	3TGT	A	0.81	0.67	0.98	0.89
YU2	Unliganded	3TGQ	A	1.16	1.05	1.25	1.46
E(93TH057)	VRC01	3NGB	A	0.81	0.67	0.98	1.07

[1] Cα-RMSDs of a set of common residues shared by all gp120s ($RMSD_{Com}$), of inner domain ($RMSD_{ID}$), of outer domain ($RMSD_{OD}$) and of bridging sheet ($RMSD_{BS}$) were calculated after each of the gp120 structures was superimposed on the VRC01-bound gp120 structure. The common set consists of 253 residues, including residues 90-120, 204-205, 215-298, 329-355, 357-392, 395, 413-469 and 473-489. The inner domain (ID) contains three segments, residues 83-118, 206-252 and 476-492. The outer domain (OD) contains three segments, residues 253-397, 410-419 and 437-475. The bridging sheet contains two regions, residues 118-205 and 421-437.

[2] For 3LQA, which has adopted a different residue numbering, and 2BF1, which is the gp120 from SIV, sequence alignment with 1G9M is used to derive the standard HXBc2 numbering.

Table S7. Isothermal titration calorimetry

Clade B_{YU2} gp120 $core_e$ (44-492 ΔV12/V3s)

gp120	Ligand	$\Delta G_{(25°C)}$	$\Delta G_{(37°C)}$	$\Delta H_{(25°C)}$	$\Delta H_{(37°C)}$	$-T\Delta S_{(25°C)}$	$-T\Delta S_{(37°C)}$
$core_e$	CD4	-11.1	-12.2	-19.5	-26.5	8.4	14.3
$core_e$	VRC01Fab	-12.3	-11.2	-26.5	-36.9	14.2	25.7
$core_e$	17b Fab	-12.6	-12.8	-16.5	-45.3	3.9	24.2

Clade B_{YU2} Full-length (FL) gp120

gp120	Ligand	$\Delta G_{(25°C)}$	$\Delta G_{(37°C)}$	$\Delta H_{(25°C)}$	$\Delta H_{(37°C)}$	$-T\Delta S_{(25°C)}$	$-T\Delta S_{(37°C)}$
YU2 FL	CD4	-10.5	-10.7	-27.0	-55.6	16.5	45.0
YU2 FL	VRC01Fab	-11.0	-11.2	-26.7	-46.9	15.7	35.7
YU2 FL	17b Fab	-10.9	-11.0	-24.0	-51.6	13.1	40.6

Clade B_{YU2} Full-length gp120 (S375W)

gp120	Ligand	$\Delta G_{(25°C)}$	$\Delta G_{(37°C)}$	$\Delta H_{(25°C)}$	$\Delta H_{(37°C)}$	$-T\Delta S_{(25°C)}$	$-T\Delta S_{(37°C)}$
YU2 FL(S375W)	CD4	-	-11.1	-	-33.9	-	22.8
YU2 FL(S375W)	VRC01Fab	-	-11.2	-	-33.9	-	22.7
YU2 FL(S375W)	17b Fab	-	-10.9	-	-37.3	-	26.4

Clade C_{1086} gp120 $core_e$ (44-492 ΔV12/V3s)

gp120	Ligand	$\Delta G_{(25°C)}$	$\Delta G_{(37°C)}$	$\Delta H_{(25°C)}$	$\Delta H_{(37°C)}$	$-T\Delta S_{(25°C)}$	$-T\Delta S_{(37°C)}$
$core_e$	CD4	-11.4	-11.7	-19.4	-32.9	8.1	21.9
$core_e$	VRC01Fab	-10.5	-10.3	-19.6	-39.8	8.3	29.6
$core_e$	17b Fab	-10.5	-11.9	-12.0	-31.6	1.05	19.8

Clade C_{1086} Full-length (FL) gp120

gp120	Ligand	$\Delta G_{(25°C)}$	$\Delta G_{(37°C)}$	$\Delta H_{(25°C)}$	$\Delta H_{(37°C)}$	$-T\Delta S_{(25°C)}$	$-T\Delta S_{(37°C)}$
1086 FL	CD4	-9.0	-11.3	-21.5	-51.6	12.5	40.3
1086 FL	VRC01Fab	-9.4	-9.8	-18.2	-47.0	8.8	37.2
1086 FL	17b Fab	-11.0	-11.1	-18.3	-52.3	7.3	41.2

Table S8. **Kinetic parameters on gp120 variants interactions with NBD-556 or sCD4**

Ligand: NBD-556 (~6.5 RU)

Analyte	K_D (μM)	% bound	R_{max}
YU2 $core_{min}$	370.2	6.0 (7.5)	185.2
YU2 $core_e$	4.1	79.1 (100)	185.2
YU2 44-492 ΔV1/2	4.7	77.5 (97.9)	185.2
YU2 44-492 ΔV3	21.8	46.9 (59.2)	185.2
YU2 Full-length	76.1	22.3 (28.1)	185.2

K_D's were calculated by fitting responses from all analytes globally and sharing the same R_{max} (185.2 RU) for gp120 variants : NBD-556 binding. Numbers in parenthesis represent % bound analytes to ligand, NBD-556, normalized to % bound YU2 $core_e$ to NBD-556.

Ligand: sCD4 (~200 RU)

Analyte	ka (1/Ms) ([1])	kd (1/s)	K_D (nM)	R_{max} (RU)
YU2 $core_{min}$	2.67×10^4 (1.64)	1.26×10^{-3}	47.0	26.0
YU2 $core_e$	4.66×10^4 (2.88)	1.87×10^{-4}	4.0	42.1
YU2 44-492 ΔV1/2	3.16×10^4 (1.95)	1.74×10^{-4}	5.5	39.9
YU2 44-492 ΔV3	4.95×10^4 (3.05)	6.07×10^{-4}	12.3	41.4
YU2 Full-length	1.62×10^4 (1.00)	8.90×10^{-4}	55.0	35.5

[1] Fold increase over full-length

Table S9. Neutralization profiles of sCD4, NBD-556, and other ligands against clade B and C reference viruses

		IC_{50} values (μM)				IC_{80} values (μM)			
		sCD4	**NBD-556**	**17b IgG**	**48d IgG**	**sCD4**	**NBD-556**	**17b IgG**	**48d IgG**
Tier 1 clade B (n=7)	**HXB2**	0.0005	4.7	0.007	0.002	0.001	19	0.043	0.224
	MN.3	0.0005	67	0.002	0.001	0.001	>100	0.009	0.002
	BaL.01	0.001	42	>0.5	>0.5	0.003	78	>0.5	>0.5
	BaL.26	0.002	39	0.132	>0.5	0.005	88	>0.5	>0.5
	ADA	0.004	17	0.037	0.064	0.013	68	>0.5	>0.5
	SF162	0.005	18	0.017	0.093	0.013	37	0.313	>0.5
	SS1196.1	0.013	13	0.026	0.126	0.052	34	0.340	>0.5
Tier 2 clade B (n=18)	**YU2**	0.004	75	>0.5	>0.5	0.010	>100	>0.5	>0.5
	89.6	0.006	11	>0.5	>0.5	0.015	26	>0.5	>0.5
	THRO.18	0.010	43	>0.5	>0.5	0.035	>100	>0.5	>0.5
	REJO.67	0.017	12	>0.5	>0.5	0.107	70	>0.5	>0.5
	JR-FL	0.017	>100	>0.5	>0.5	0.038	>100	>0.5	>0.5
	QH0692.42	0.034	37	>0.5	>0.5	0.098	99	>0.5	>0.5
	6535.3	0.050	12	>0.5	>0.5	0.383	33	>0.5	>0.5
	RHPA.7	0.050	>100	>0.5	>0.5	0.249	>100	>0.5	>0.5
	7165.18	0.077	11	>0.5	>0.5	0.342	34	>0.5	>0.5
	6101.10	0.109	11	>0.5	>0.5	0.183	19	>0.5	>0.5
	PVO.4	0.151	80	>0.5	>0.5	0.457	>100	>0.5	>0.5
	WITO.33	0.185	43	>0.5	>0.5	0.506	>100	>0.5	>0.5
	SC422.8	0.205	>100	>0.5	>0.5	>1.5	>100	>0.5	>0.5
	TRJO.58	0.362	87	>0.5	>0.5	0.632	>100	>0.5	>0.5
	CAAN.A2	0.550	8.3	>0.5	>0.5	>1.5	18	>0.5	>0.5
	AC10.29	0.554	100	>0.5	>0.5	1.440	>100	>0.5	>0.5
	TRO.11	0.599	>100	>0.5	>0.5	>1.5	>100	>0.5	>0.5
	BG1168.1	0.716	14	>0.5	>0.5	>1.5	30	>0.5	>0.5
	breadth n=25	**100%**	**84%**	**24%**	**20%**	**84%**	**56%**	**16%**	**8%**
	geometric mean	**0.028**	**25**	**0.018**	**0.012**	**0.049**	**40**	**0.080**	**0.021**
Tier 1 clade C	**MW965.26**	0.005	53	0.001	0.305	0.069	>100	0.002	>0.5
Tier 2 clade C (n=16)	**ZM109.4**	0.005	98	>0.5	>0.5	0.020	>100	>0.5	>0.5
	Du123.6	0.006	>100	>0.5	>0.5	0.016	>100	>0.5	>0.5
	Du172.17	0.041	65	>0.5	>0.5	0.116	>100	>0.5	>0.5
	Du151.2	0.058	>100	>0.5	>0.5	0.142	>100	>0.5	>0.5
	CAP210.E8	0.078	>100	>0.5	>0.5	0.342	>100	>0.5	>0.5
	ZM233.6	0.118	42	>0.5	>0.5	0.322	>100	>0.5	>0.5
	CAP244.D3	0.134	92	>0.5	>0.5	0.395	>100	>0.5	>0.5
	ZM197.7	0.153	74	>0.5	>0.5	0.723	>100	>0.5	>0.5
	ZM53.12	0.163	>100	>0.5	>0.5	0.455	>100	>0.5	>0.5
	ZM214.15	0.169	>100	>0.5	>0.5	>1.5	>100	>0.5	>0.5
	ZM135.10a	0.186	>100	>0.5	>0.5	0.814	>100	>0.5	>0.5
	ZM249.1	0.190	95	>0.5	>0.5	0.693	>100	>0.5	>0.5
	Du422.1	0.221	>100	>0.5	>0.5	0.646	>100	>0.5	>0.5
	Du156.12	0.450	76	>0.5	>0.5	1.42	>100	>0.5	>0.5
	ZM106.9	1.114	>100	>0.5	>0.5	>1.5	>100	>0.5	>0.5
	CAP45.G3	1.5	>100	>0.5	>0.5	>1.5	>100	>0.5	>0.5
	Breadth n=17	**100%**	**47%**	**6%**	**6%**	**82%**	**0%**	**6%**	**0%**
	Geometric mean	**0.101**	**72**	**0.001**	**0.305**	**0.244**	**>100**	**0.002**	**>0.5**
SIV control	**SIVmac251.30.SIV**	0.060	88	>0.5	>0.5	0.847	>100	>0.5	>0.5
	MuLV.SG3	>1.5	>100	>0.5	>0.5	>1.5	>100	>0.5	>0.5