Supplementary Information for

Mutational fitness landscapes reveal genetic and structural improvement pathways for an HIV-1 broadly neutralizing antibody

Bharat Madan[1,7], Baoshan Zhang[2,7], Kai Xu[2], Cara W. Chao[2], Sijy O'Dell[2], Jacy R. Wolfe[1], Gwo-Yu Chuang[2], Ahmed S. Fahad[1], Hui Geng[2], Rui Kong[2], Mark K. Louder[2], Thuy Duong Nguyen[1], Reda Rawi[2], Arne Schön[3], Zizhang Sheng[4], Rajani Nimrania[1], Yiran Wang[2], Tongqing Zhou[2], Bob C. Lin[2], Nicole A. Doria-Rose[2], Lawrence Shapiro[2,4,5], Peter D. Kwong[2,4] and Brandon J. DeKosky[1,6]*

[1]Department of Pharmaceutical Chemistry, The University of Kansas, Lawrence, KS, 66045, USA
[2]Vaccine Research Center, National Institute of Allergy and Infectious Diseases, Bethesda, MD, 20814, USA
[3]Department of Biology, John Hopkins University, Baltimore, MD, 21218, USA
[4]Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY, 10027, USA
[5]Aaron Diamond AIDS Research Center, Columbia University Irving Medical Center, New York, NY, 10032, USA
[6]Department of Chemical Engineering, The University of Kansas, Lawrence, KS, 66045, USA
[7]These authors contributed equally.

*Corresponding author. Email: dekosky@ku.edu

This PDF file includes:

SI Methods
Figures S1 to S14
Tables S1 to S9

Other supplementary materials for this manuscript include the following:

Dataset_S01 (208 virus panel neutralization data)
Dataset_S02 (read counts for vFP16.02 sort data)

SI Methods

Site Saturation Mutagenesis Library Construction for vFP16.02

The VH and VL genes in the pCT_VHVL-Kappa1 (1) vector are transcribed from “+” and “-” strands, respectively via a galactose inducible bidirectional promoter Gal1/Gal10 along with CH1 (human IgG1 isotype) and CK (human Kappa) constant regions at the C-terminus of VH and VL, respectively. The C-terminus of VL also has an expression tag FLAG-tag. vFP16.02 VH and VL genes were cloned into the yeast display vector pCT_VHVL-Kappa1 using multi-fragment infusion (Takara Bio USA, Inc.; Mountain View, CA).

The cloned plasmid was used as the template for generating VH and VL site saturation mutagenesis libraries where each codon in the sequence is replaced with an NNK (N=A/T/G/C and K=G/T) or MNN (M=C/A and N=A/T/G/C) codon for heavy and light SSM libraries, respectively, that encode for all amino acids. The degenerate oligos for the mutagenesis were synthesized in a 96-well plate format (Integrated DNA Technologies, Coralville, IA). The pCT_vFP16.02 plasmid was digested with AscI and NotI and fragment containing heavy chain variable/bi-promoter/light chain variable regions were ligated into the pET_conNK plasmid (kind gift of Timothy Whitehead, University of Colorado (2)), as a starting source for generating SSM libraries.

Briefly, the wild-type backbone of template plasmid was nicked and degraded with restriction enzymes. The mutant strand was synthesized with a mixture of degenerate primers to introduce point mutations into the newly synthesized DNA strand. The opposite wild-type template strand was nicked and degraded with restriction enzymes and the second mutant strand was synthesized using a universal primer via another round of PCR amplification. The PCR product was digested with DpnI to remove any residual wild-type template and was used to transform MegaX DH10B $T1^R$ electrocompetent cells (ThermoFisher Scientific). Following transformation the cells were plated on a 255×255 mm 2xYT (Yeast Extract Tryptone Medium) (Fisher Scientific) agar plates. After overnight incubation of the electroporation plates at 37°C the plates were scraped off to

collect in bacterial colonies in Terrific Broth (ThermoFisher Scientific). 1 mL of this mixture was maxiprepped (2) and resulting DNA libraries were used to amplify the insert for yeast transformation described below.

Transformation of SSM Libraries into Yeast

The AWY101 yeast strain was used for yeast library construction and screening (1, 3). Untransformed yeast cells were maintained in YPD medium (20 g/l dextrose, 20 g/l peptone, and 10 g/l yeast extract). VH and VL libraries were PCR amplified using the following conditions: 95°C for 2 min; 15 cycles of 98°C for 20 sec, 55°C for 15 sec, 72°C for 15 sec and final extension at 72°C for 2 min (1). PCR reactions for each library were cleaned using a PCR cleanup kit and linear DNAs were used to transform yeast in combination with linear pCT_VHVL-Kappa1 plasmid pre-digested with AscI and NotI restriction enzymes to generate the complete plasmid containing the entire Fab expression cassette in the AW101 yeast strain using the LiAC/DTT transformation method (4). Yeast cells were grown in 250 mL of SDCAA medium after library transformation for 16-20 hrs (TEKnova, Hollister, CA; 20 g/l dextrose, 6.7 g/l yeast nitrogen base, 5 g/l casamino acids, 8.56 g/l $NaH_2PO_4.H_2O$, and 10.2 g/l $Na_2HPO_4.7H_2O$) and passaged twice in SDCAA to maintain a single copy of the plasmid per yeast cell. We referred to these transformed libraries after two passages in yeast as pre-sort or input libraries.

Bioinformatic Analysis

ERs for all single variants of the VH and VL SSM libraries in three affinity groups (*i.e.*, low, med, and high) were calculated using Eq. 1. Single variants were selected for experimental validation based on ER in high affinity groups in Round 2 and Round 3. We rejected any sequences in any library across all rounds of sorting that have read count <2 for ER based selection of clones with low, med and high affinity.

While selecting for high affinity variants, we verified that a particular variant did not enrich in low or medium affinity groups. Binning of mutations as low, med or high was performed based on prevalence of a mutation in each of the three affinity groups. Each mutation was assigned to an affinity group with the highest observed prevalence. i.e., the ER_{max} of each clone across high vs. medium vs. low affinity sorted populations.

Heat maps were generated using the imshow() (5) function of the matplotlib plotting library in Python (v3.6). Scatter plots were generated using the ggscatter() function of the ggpubr library in R (v3.6.2) (6). Pearson correlations for the scatter plots were calculated using stat_cor function of the ggpubr library in R (v3.6.2). A two-tailed Student's t-test was used to determine statistical significance ($p < 0.05$) for five analyses (**Figs 5B, 6D, S8, S11, S12**).

We built gene-specific substitution profiles for mouse IGHV1-15 and IGKV1-117 genes to infer mutability of individual mutations. Briefly, we simulated the somatic hypermutation process starting from the germline V-genes with mutation frequency and substitution preference at each position weighted using the S5F model (7), which predicts mutability and substitution preference of AID. For comparison to the vFP16.02 antibody, we generated sequences having identical SHM levels in both VH and VL, which are 16 and 13 nucleotide mutations respectively. Sequences with stop codons were removed and a total of 10,000 sequences were generated for each gene. We then calculated mutation frequency of each AA at each V position using our bioinformatics script (8).

Antibody Expression and Characterization

VH and VL genes of vFP16.02 variants were generated by site-directed point mutation on vFP16.02 antibody genes (Gene Universal, Newark, Delaware). When used in fab production, heavy chain plasmids were constructed carrying a HRV3C cleavage site inserted in the hinge region. VH and VL plasmids (1:1 DNA ratio) were co-transfected into Expi293F cells using Turbo293 transfection reagent (Speed Biosystems, Gaithersburg, Maryland) following manufacturer's protocols. After transfection, cells were incubated in shaker incubators at 120 rpm, 37°C, 9%

CO_2 overnight. On the next day, one tenth culture volume of AbBooster medium (ABI scientific Inc, STERLING, Virginia) was added to each flask of transfected cells. Cell cultures were incubated at 120 rpm, 33°C, 9% CO_2 for an additional 5 days. Cell culture supernatant was harvested at 6 days post transfection. IgGs were purified from the supernatants using protein A chromatography. Fabs were obtained through on column HRV3C cleavage. The fabs were further purified by size exclusion chromatograph (SEC) in a Superdex 200 column (GE) with a PBS buffer (for affinity analysis) or buffer containing 5mM HEPES, pH 7.5, 150mM NaCl (for protein crystallization).

Affinity Assays

Octet for soluble FP

A fortéBio Octet HTX was used to measure binding of soluble FP to the vFP16.02 mouse antibody variants. Assays were performed in tilted black 384-well plates (Geiger Bio-One) with 1x kinetic buffer (purchased from fortéBio) with agitation set to 1,000 rpm. Mouse antibody IgG at 20 µg/mL were immobilized using Anti-Mouse IgG Fc biosensors for 300 s. To measure binding of vFP16.02 IgG variants to soluble FP-1M6T, the immobilized IgG were dipped into solutions of FP at a two-fold serial dilution with a starting concentration of 100 nM for 60 s. Reference sensors were loaded with mouse IgG and dipped into 1x kinetic buffer. Baseline drift was corrected for through parallel subtraction of reference sensors from sample measurements and fitted globally using a 1:1 model.

SPR for trimer affinity

Binding affinities and kinetics of vFP16.02 mouse antibodies to HIV-1 Env were assessed on a Biacore T200 (GE Healthcare) in HBS-EP+ buffer (10 mM HEPES, pH7.4, 3 mM EDTA, 150 nM NaCl, and 0.05% surfactant P-20) at 25° C. To prepare the chip, 2G12 IgG was immobilized onto a CM5 chip by amine coupling to 8,000-10,000 response units (RU). BG505 DS-SOSIP at 250nM was then captured to ~400 RU in the sample flow cell on the 2G12 sensor chip. To measure affinity to antibody Fabs, two-fold serial dilution of mouse vFP16.02 Fabs were

injected into the sample and reference cells for 180 s followed by 300 s of dissociation at a flow rate of 30uL/min. The starting concentration of the Fabs were varied between 25 nM and 5 nM depending on their affinity to the BG505 DS-SOSIP trimer. The chip surface was regenerated using 3M $MgCl_2$ at a flow rate of 50 µL/min for 30 s. Reference sensorgrams were obtained by flowing equal volumes of 1x HBS-EP+ buffer instead of antibody Fab solution. Sensorgrams from the concentration series were corrected to the blank reference curves and fitted globally using a 1:1 Langmuir model through the BIAevaluation software.

X-ray Crystallization and Structural Analysis

Crystals of S48K Fab complexed with FP8 were obtained from 0.1 M Tris-HCl pH8.5 and 30% v/v PEG 400 and grew to dimensions of 200 um x 200 um x 100 um. Crystals of F60P Fab complexed with FP8v1 were obtained from 9.9% v/v Isopropanol, 0.1 M Tris-HCl pH8.5, and 9.9% v/v PEG 3350 and grew to dimensions of 200 um x 200 um x 150 um. Crystals were cryoprotected in 25% glycerol and flash-frozen in liquid nitrogen. Data were collected at 100 K and a wavelength of 1.00 Å at the SER-CAT beamline ID-22 (Advanced Photon Source, Argonne National Laboratory). Diffraction data were processed with HKL2000 (9) . Structure solution was obtained by molecular replacement with Phaser (10) using homologous vFP16.02 Fab:FP complex structure (PDB ID: 6CDO) as search model. Model building was carried out with Coot (11), followed by refinement with Phenix (12). Structure parameter were validated by MOLPROBITY (13) during refinement. Ramachandran statistical analysis indicated that the final structures contained no more than 0.2% disallowed residues. Data collection and refinement statistics are shown in Table S9.

Isothermal Titration Calorimetry

The results are expressed per mole of Fab and the stoichiometry, *N*, denotes the number of binding sites per mole of trimer. The heat associated with binding was obtained from the integral of the calorimetric signal after adjustment of the

heat of dilution. The enthalpy change, ΔH, the association constant, K_a (the dissociation constant, K_d =1/K_a) and the stoichiometry, N, were obtained by nonlinear least square analysis of the data to a single-site binding model using Origin with a fitting function made inhouse. Gibbs energy, ΔG, was calculated from the binding affinity using $\Delta G = -RT\ln K_a$, (R = 1.987 cal/(K × mol) and T is the absolute temperature in kelvin). The entropy contribution to Gibbs energy, $-T\Delta S$, was calculated from the relation $\Delta G = \Delta H - T\Delta S$.

Molecular Dynamics Simulation

The modeled Fab structures for VL-S48K and VL-F60P were solvated in a 15Å padding water box and neutralized by the addition of NaCl at a concentration of 150 mM. The final systems were composed of about 47000 atoms and measured 97 x 69 x 77 $Å^3$.

The MDs were performed using NAMD2.13 engine (14), with the CHARMM36 force field(15, 16). TIP3P water parameterization was utilized to describe the water molecules(17). The periodic electrostatic interactions were computed using particle-mesh Ewald (PME) summation with a grid spacing smaller than 1 Å. Constant temperature was imposed by using Langevin dynamics with a damping coefficient of 5.0 ps. The system was first minimized by 20000 conjugate gradient steps and then equilibrated by using a linear temperature gradient, which heated up the system from 100 to 310 K in 2 ns. An additional 10 ns were done before removing all restraints. The length of all bonds involving hydrogen atoms was constrained with the RATTLE algorithm, thus allowing a time step of 2 fs. Unrestrained molecular dynamics were performed up to 500 ns.

ABangle (18) was used to characterize the VH:VL interface angle. Statistical significance was determined using Mann-Whitney U test.

Supplementary Figures

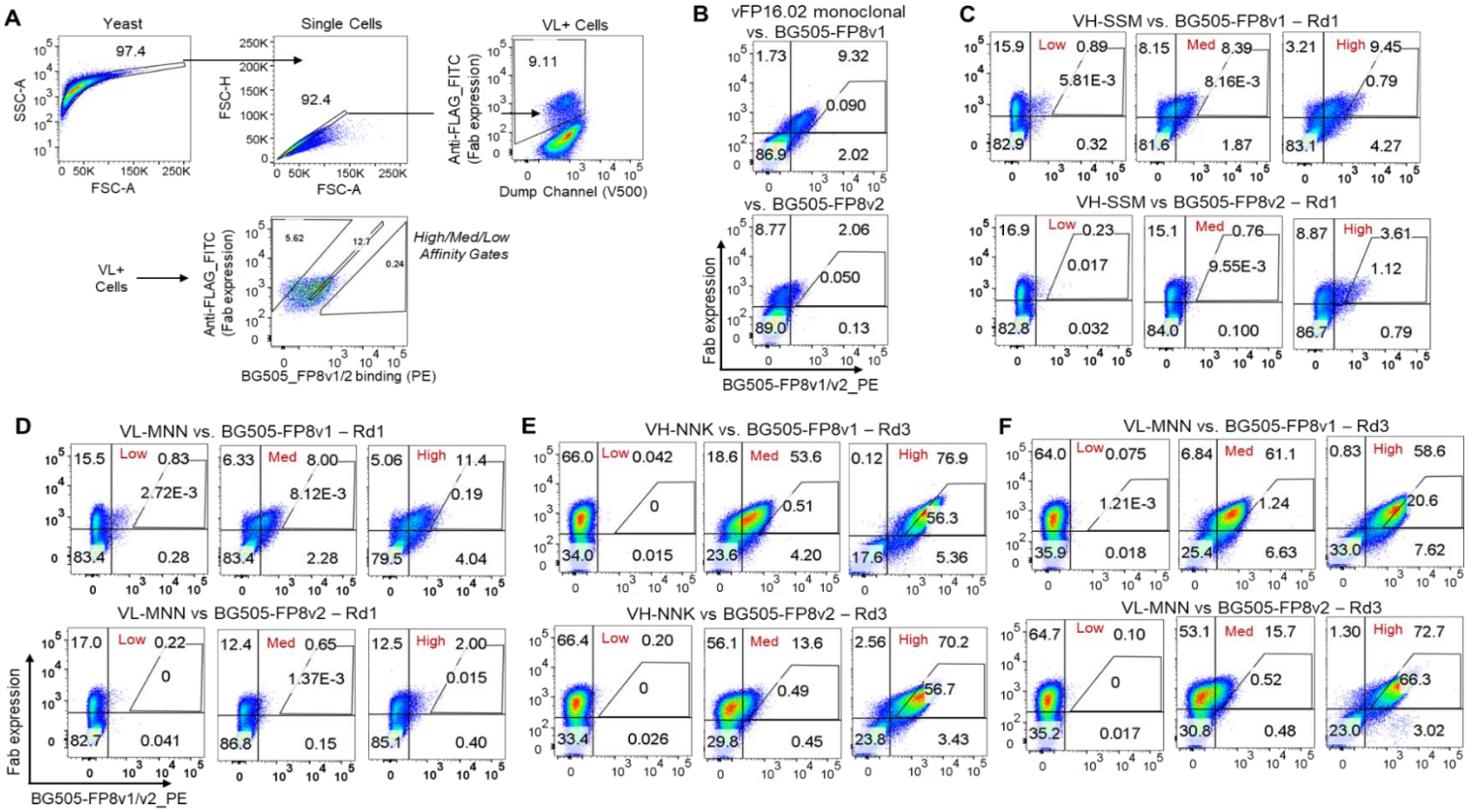

Figure S1. Affinity based sorting of vFP16.02 heavy and light chain SSM libraries against BG505-FP8v1 and -v2. (A) FACS gating strategy for affinity-based sorting of SSM libraries. **(B)** FACS analysis for vFP16.02 monoclonal fab stained against BG505-FP8v1 at 70 nM (top panel) and BG505-FP8v2 (100 nM). **(C)** and **(D)** FACS analysis of Round 1 sorted Low/Med/High populations showing phenotypic shifts when comparing between the Low, Med, and High gated populations. **(E)** and **(F)** FACS analysis of Round 3 sorted and enriched populations showing a higher degree of phenotypic population shifts as compared to Round 1. Upper panels in **(B)** to **(F)** show FACS plots for libraries stained against BG505-FP8v1 (70 nM), and bottom panels show libraries stained against BG505-FP8v2 (100 nM).

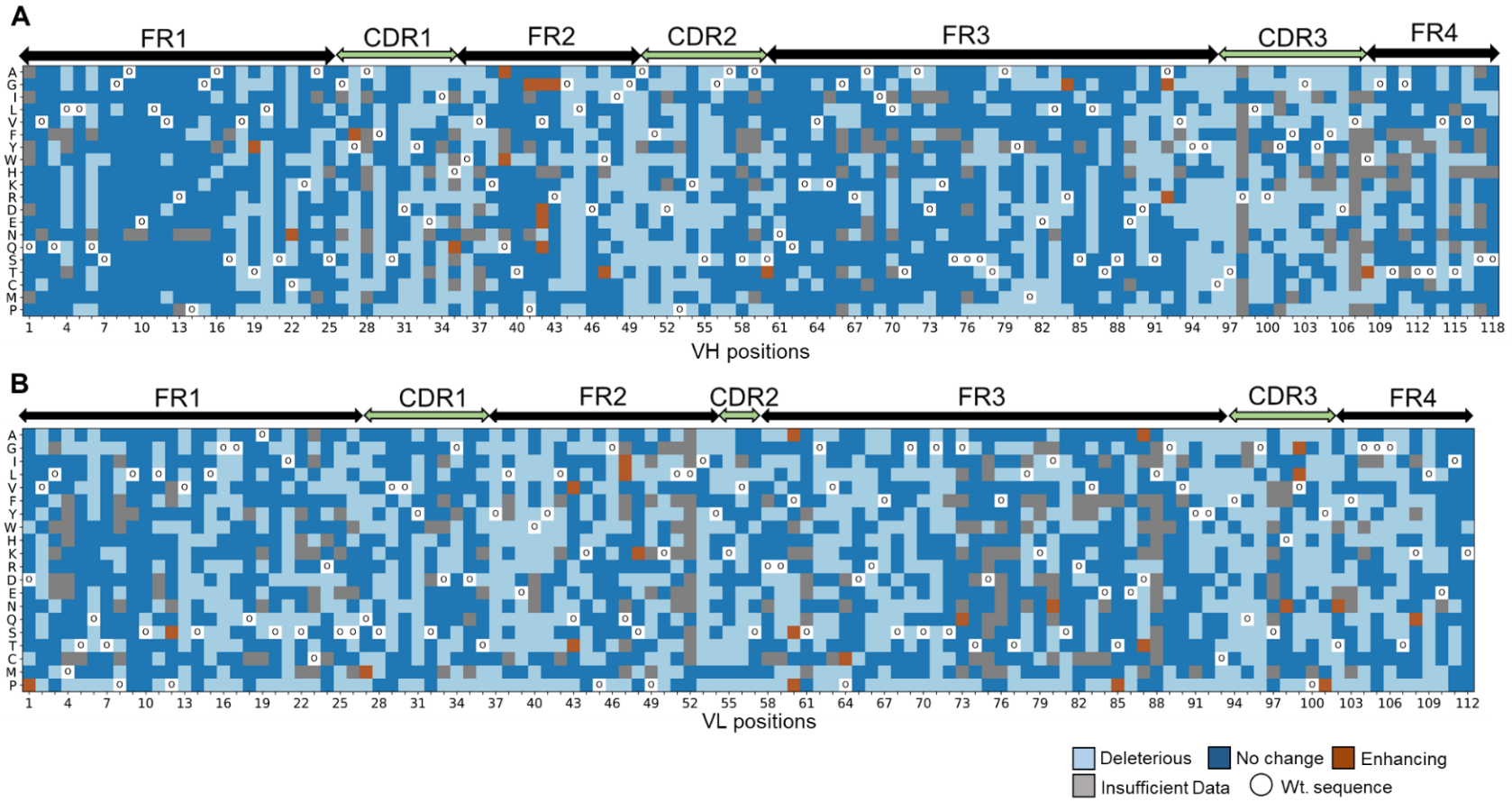

Figure S2. Heat maps showing the functional impact of single mutations in the vFP16.02 variable regions. **(A)** Heavy chain and **(B)** Light chain variable regions sorted against BG505-FP8v2.

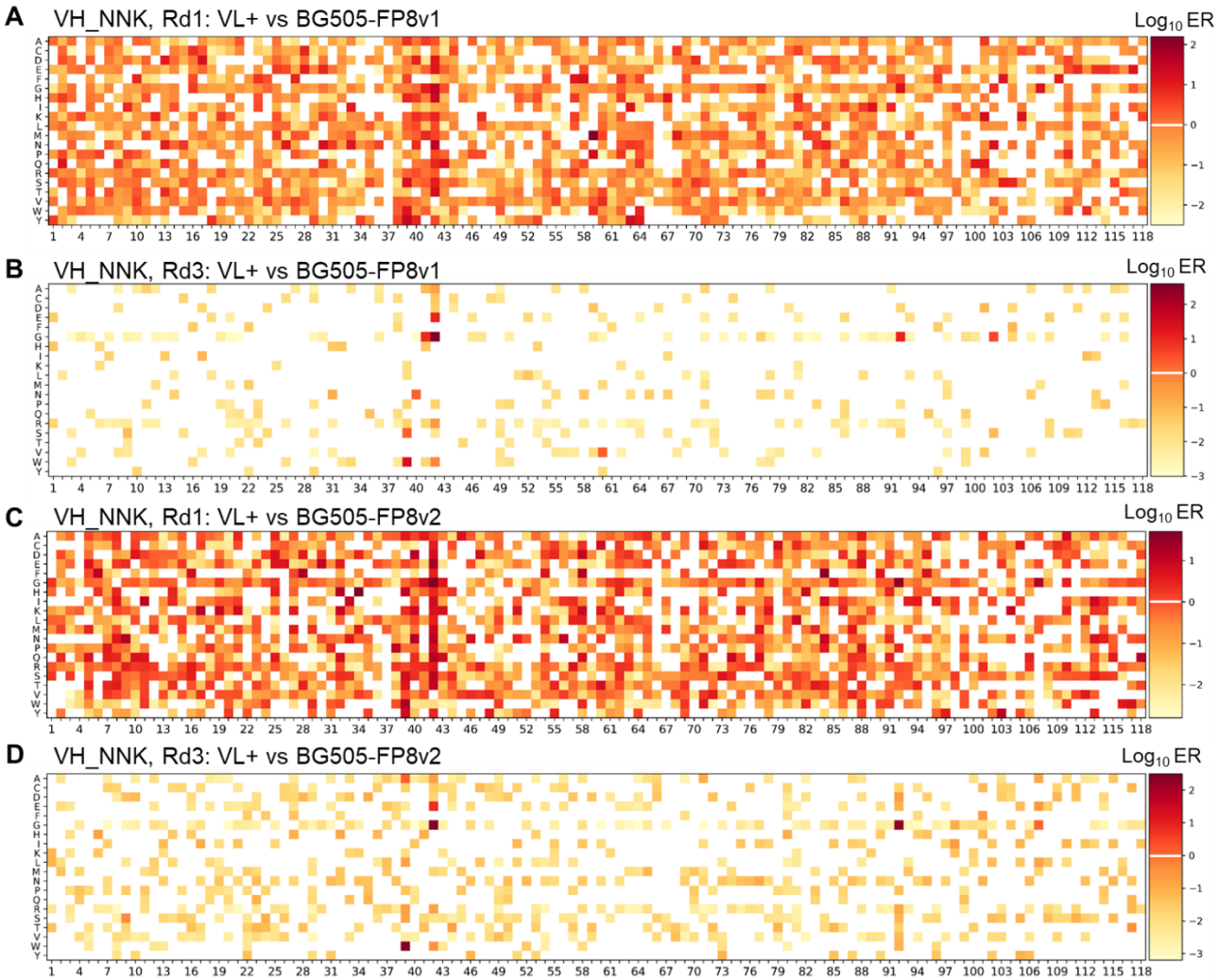

Figure S3. Mutational landscape of high affinity gate population of VH_NNK libraries against BG505-FP8v1 and v2, shown as ER heat maps. (A) VH_NNK library against BG505-FP8v1 after the first round of enrichment, and **(B)** after the third round of enrichment. **(C)** VH_NNK library against BG505-FP8v2 after the first round of enrichment, and **(D)** after the third round of enrichment.

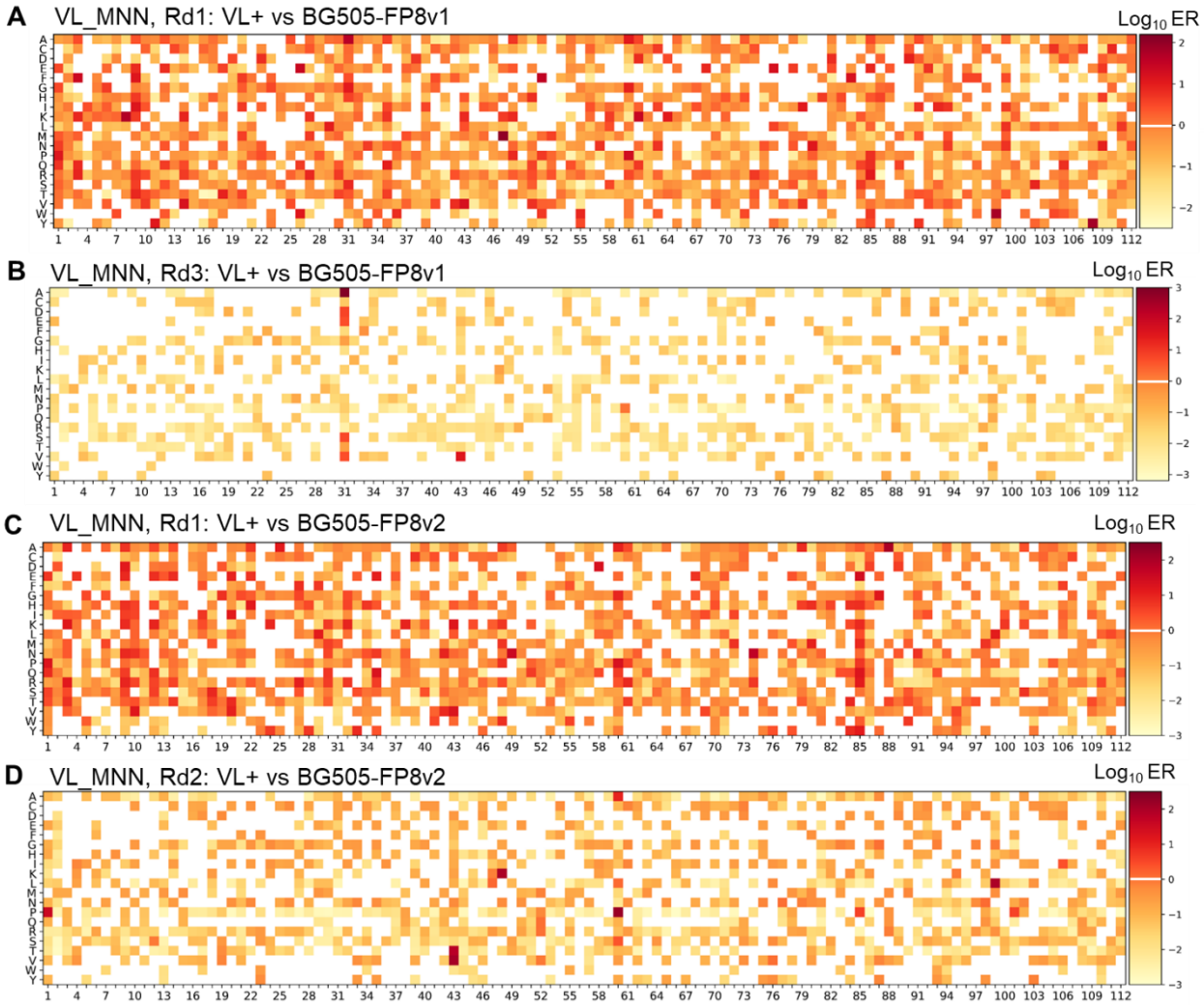

Figure S4. Mutational landscape of high affinity gate population of VL-MNN libraries against BG505-FP8v1 and v2 shown as ER heat maps. (A) VL-MNN library against BG505-FP8v1 after the first round of enrichment, and **(B)** after the third round of enrichment. **(C)** VL-MNN library against BG505-FP8v2 after the first round of enrichment, and **(D)** after the third round of enrichment.

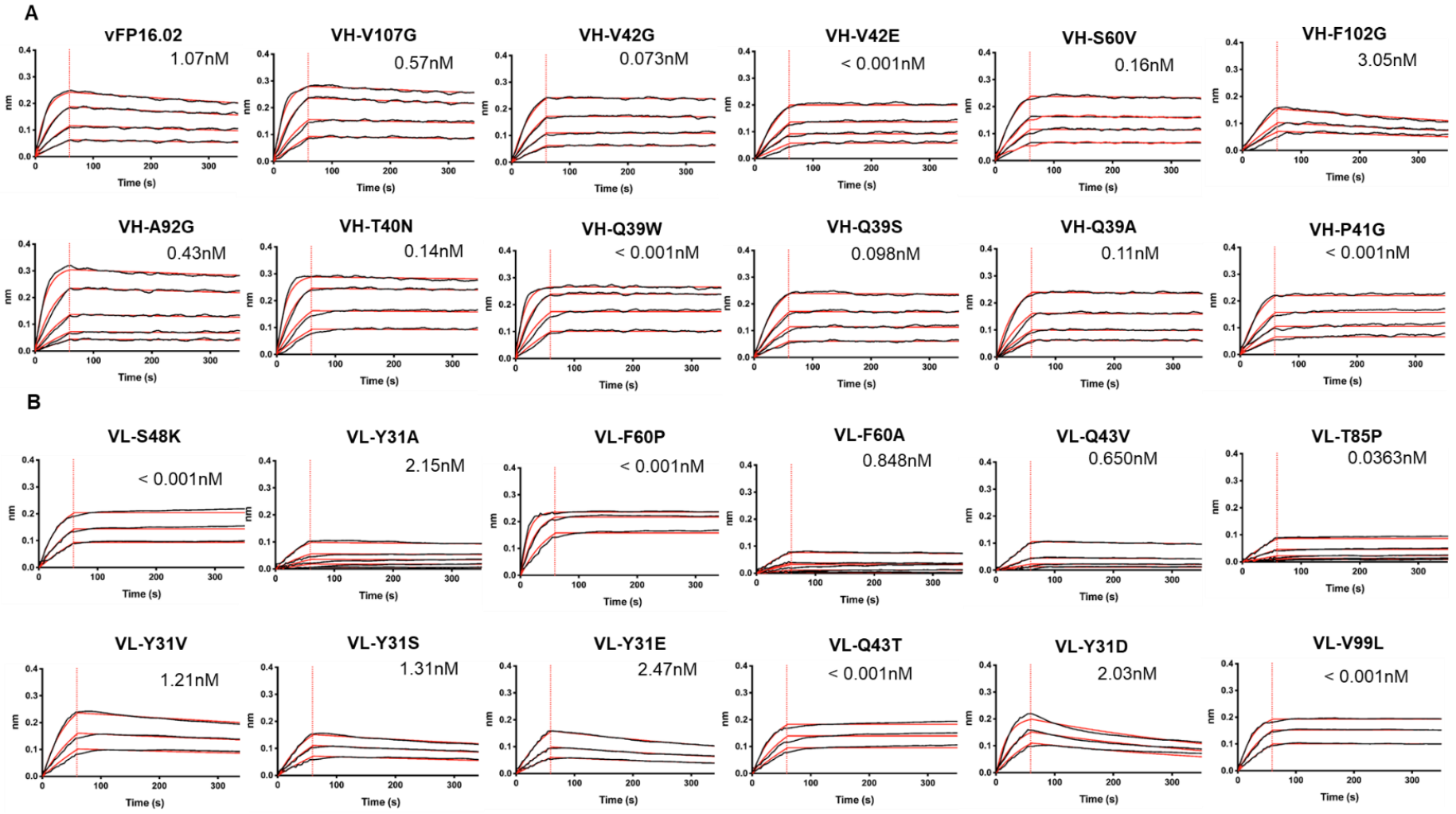

Figure S5. Bio-layer interferometry response curves for vFP16.02 single mutational variants against fusion peptide. (A) Heavy chain mutation variants, and **(B)** light chain mutation variants.

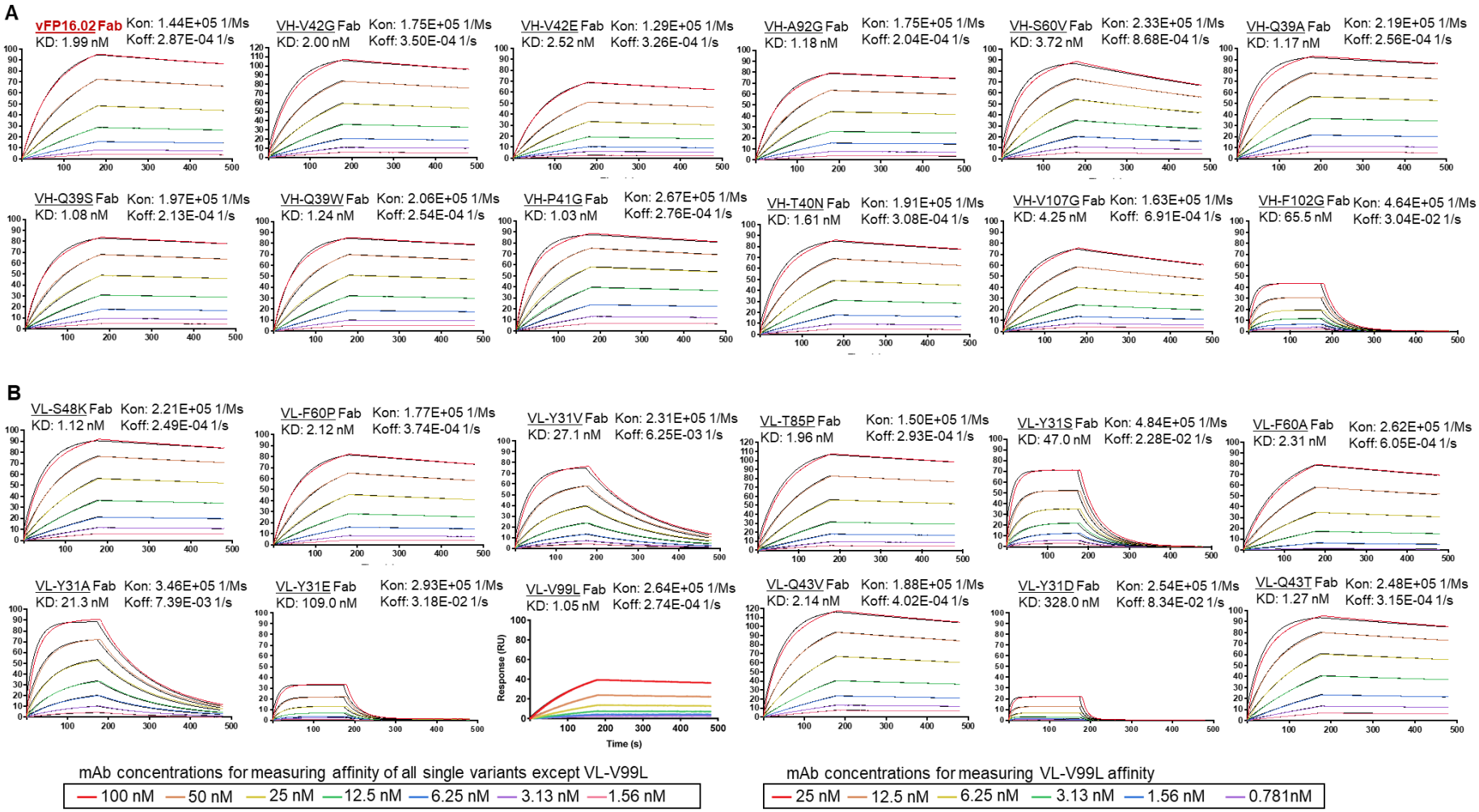

Figure S6. Surface plasmon resonance response curves for vFP16.02 single mutational variants against the BG505-FP8v1 DS.SOSIP HIV-1 Env trimer. (A) Heavy chain mutation variants, and **(B)** light chain mutation variants.

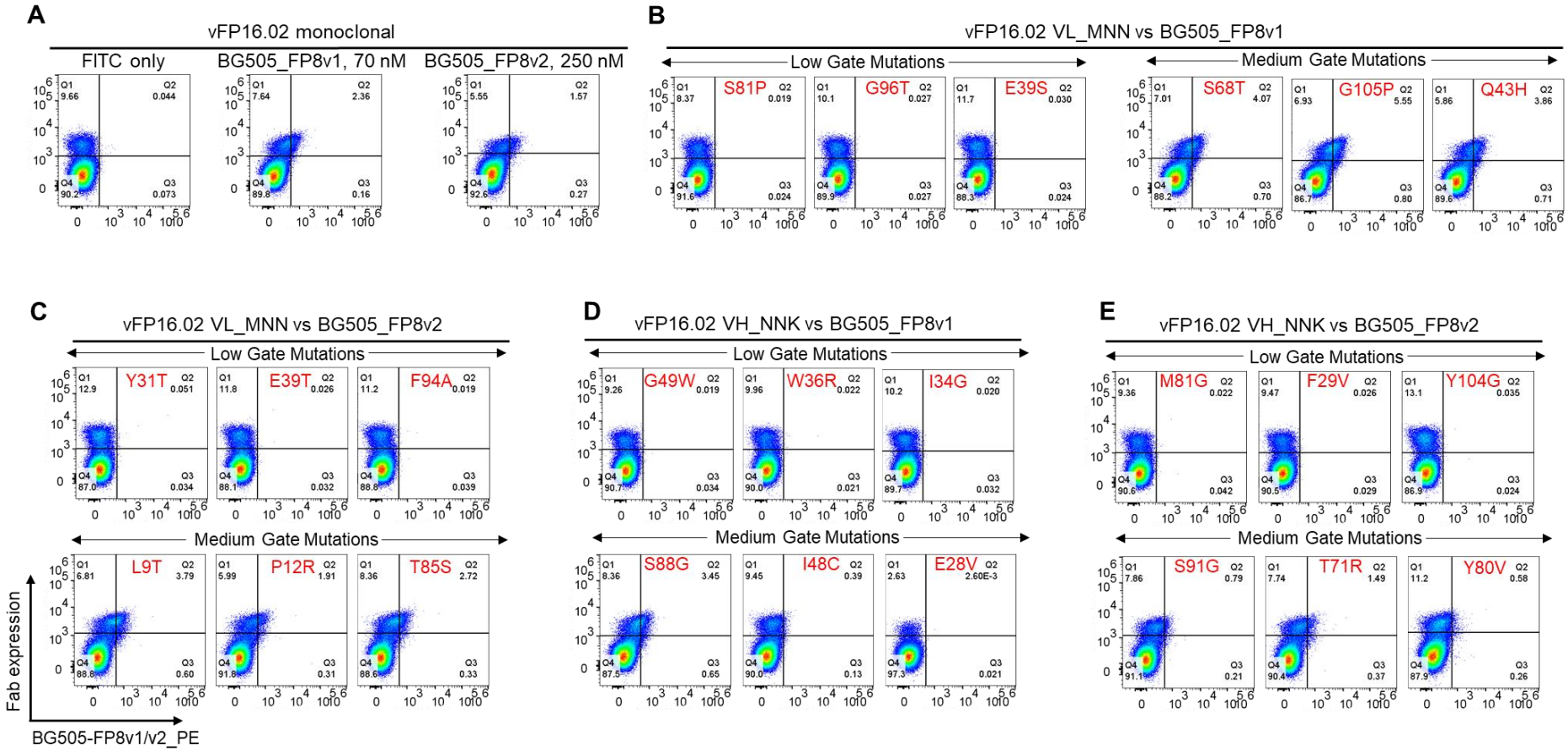

Figure S7. Evaluation of single variants enriched in low and medium gates for binding to BG505_FP8v1 and BG505_FP8v2. **(A)** FACS analyses of vFP16.02 monoclonal Fab stained against BG505_FP8v1 (70 nM) and BG505_FP8v2 (250 nM). Single variants enriched in the low and medium affinity sorts were selected from the VL_MNN library sorted against **(B)** BG505_FP8v1 and **(C)** BG505_FP8v2. Single variants enriched in the low- and medium-affinity sorts were selected from the VH_NNK library sorted against **(D)** BG505_FP8v1 and **(E)** BG505_FP8v2.

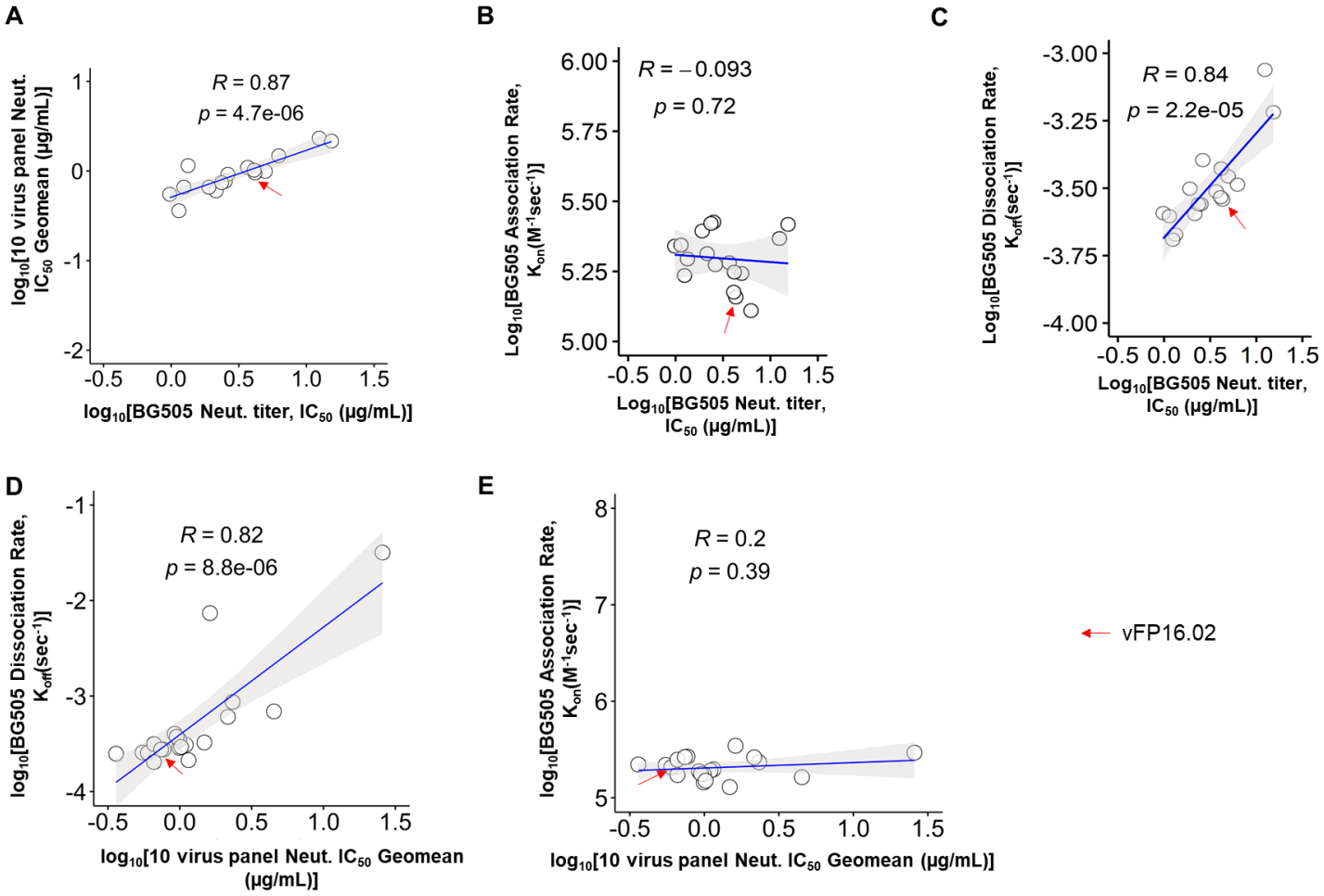

Figure S8. Correlations between biochemical parameters of vFP16.02 single variants. Correlations were plotted for variants when assessed neutralization potencies at a maximum IC_{50} level of 50 µg/mL. Shaded areas indicate 95% confidence limits.

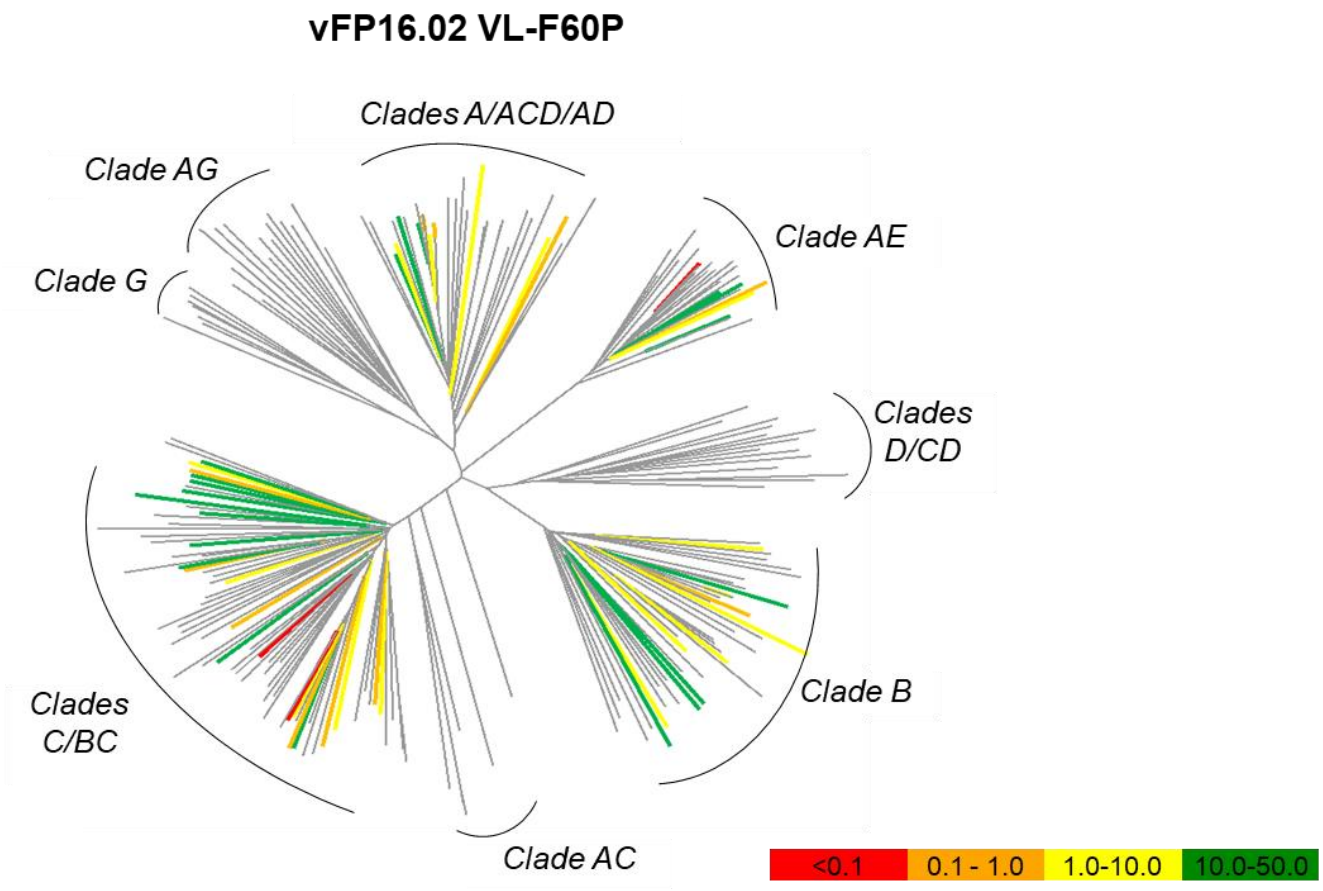

Figure S9. Neutralization dendrograms for vFP16.02 VL-F60P tested on 208-isolate panel. Also, see neutralization dendrogram for vFP16.02 in Fig. 3 in (19).

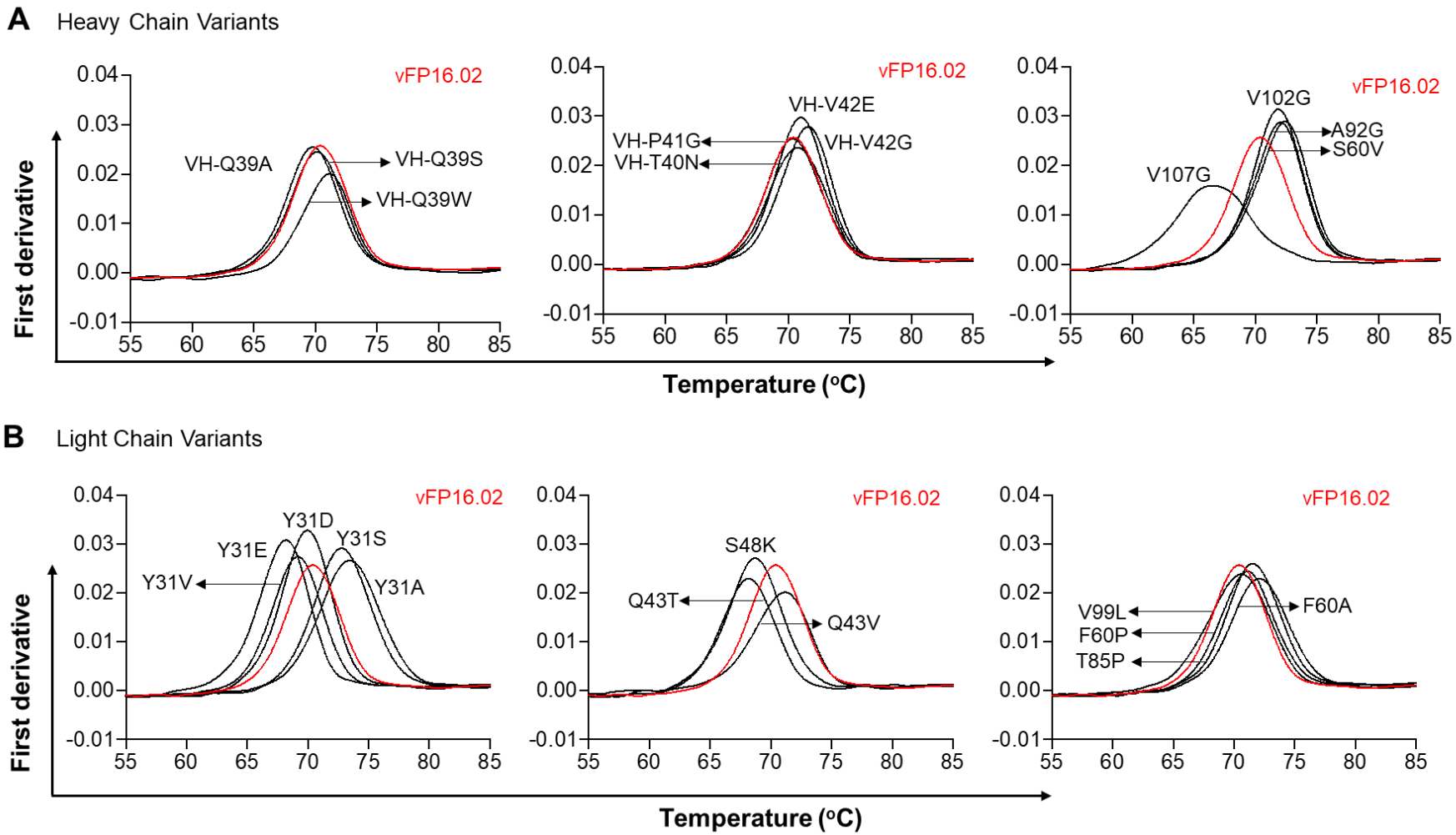

Figure S10. Effect of single vFP16.02 mutations on protein melting curves. nanoDSF measurement of T_m values of vFP16.02 variants were grouped by the locations of mutations in **(A)** heavy chains, and **(B)** light chains. vFP16.02 is shown in red, other curves are indicated by their point mutations and placed according to T_m values.

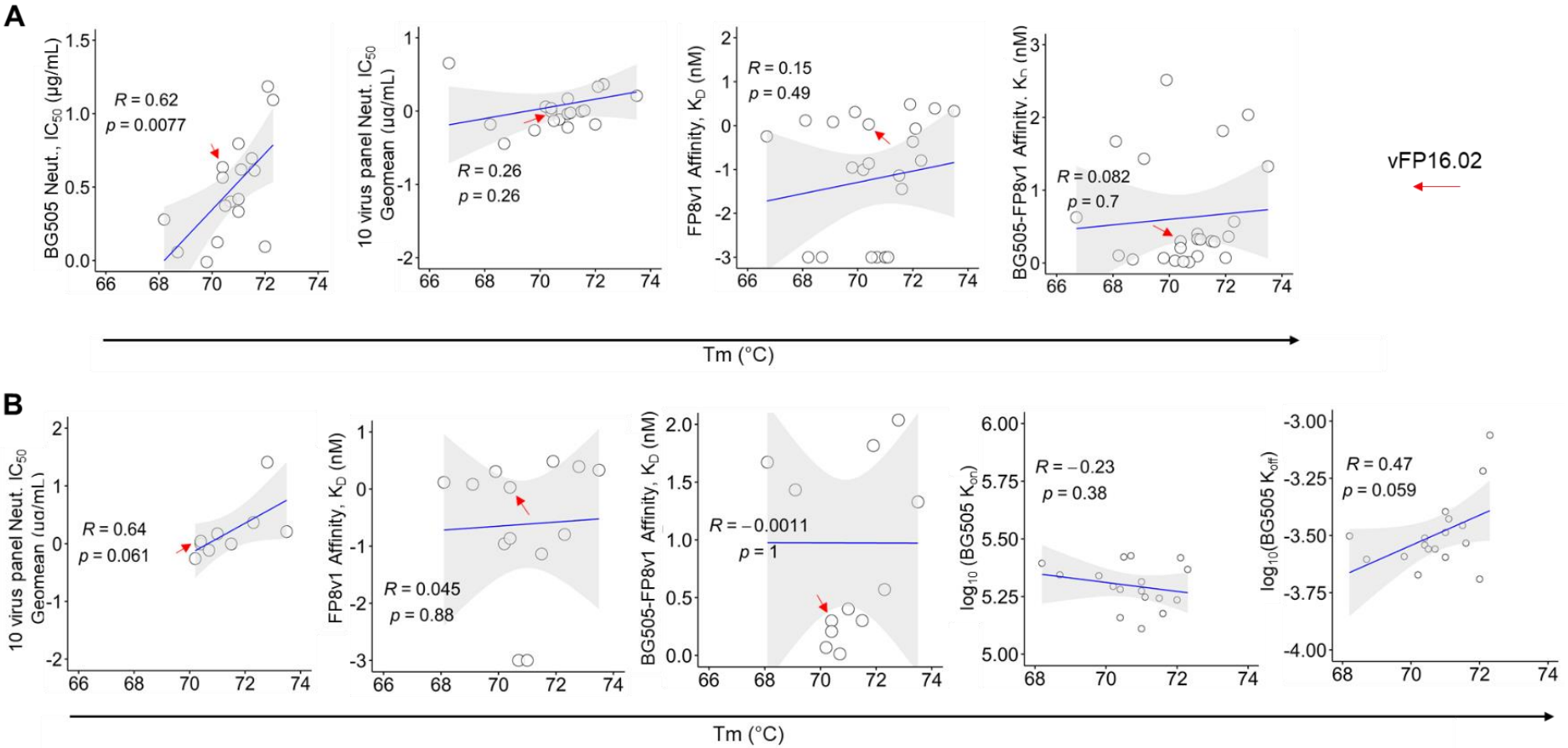

Figure S11. Correlation plots for single variants between melting temperatures and soluble FP8v1 affinity, BG505-FPv1 HIV-1 trimer Env, BG505 neutralization IC50 titers and neutralization IC50 geomean on a 10-virus panel. (A) Correlation plots for all single variants selected against BG505_FP8v1 and v2. **(B)** Correlation plots for single variants selected exclusively against BG505-FP8v1 during yeast surface display screening. Variants with neutralization IC_{50} >50 µg/mL were excluded from these analyses. Shaded areas indicate 95% confidence limits.

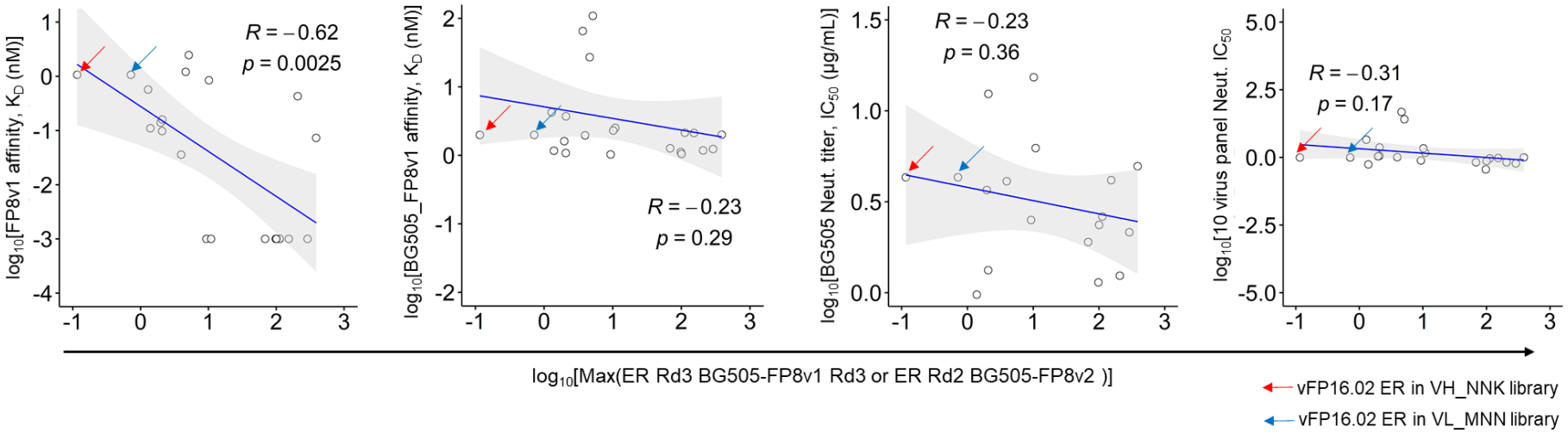

Figure S12. Correlations between functional parameters of vFP16.02 selected single variants and yeast display library enrichment ratios (ER). Correlations for all single variants selected against BG505_FP8v1 and v2 during yeast surface display screening. Different library sorts showed different magnitudes of enrichment ratios (e.g. heavy vs. light, BG505_FP8v1 vs. BG505_FP8v2), which complicates ER-based comparisons. Shaded areas indicate 95% confidence limits.

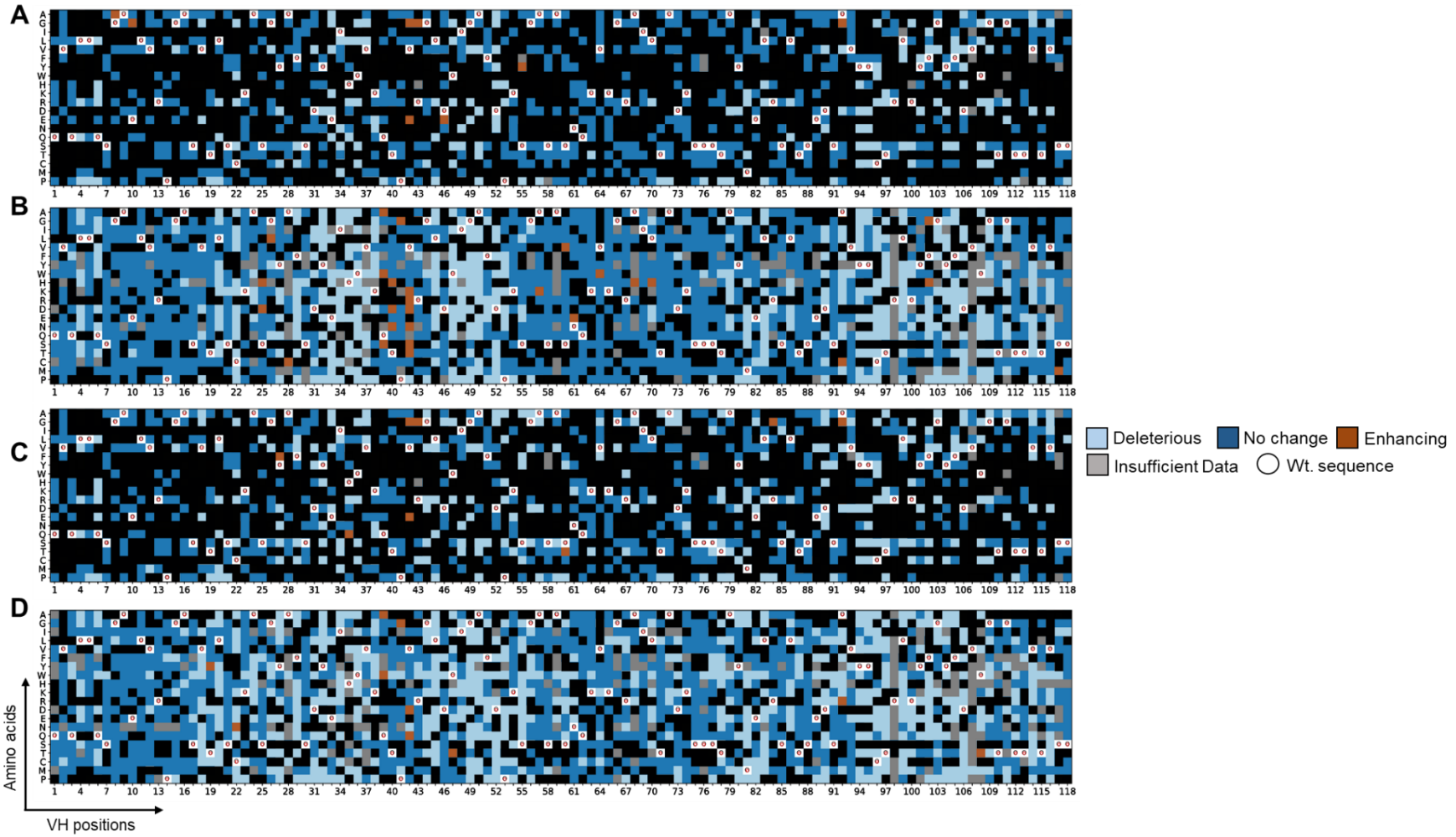

Figure S13. Heat maps highlighting functional landscapes for single vs. double/triple DNA base substitutions needed for VH single amino acid variants. (A) Single nucleotide substitutions (with multiple nucleotide substitutions blackened-out) for VH-NNK libraries against BG505-FP8v1. **(B)** Heat map for multiple (≥2) nucleotide substitutions (single nucleotide substitutions blackened-out) for VH-NNK libraries against BG505-FP8v1. **(C)** Single nucleotide substitutions (with multiple nucleotide substitutions blackened-out) for VH-NNK libraries against BG505-FP8v2. **(D)** Heat map for multiple (≥2) nucleotide substitutions (single nucleotide substitutions blackened-out) for VH-NNK libraries against BG505-FP8v2.

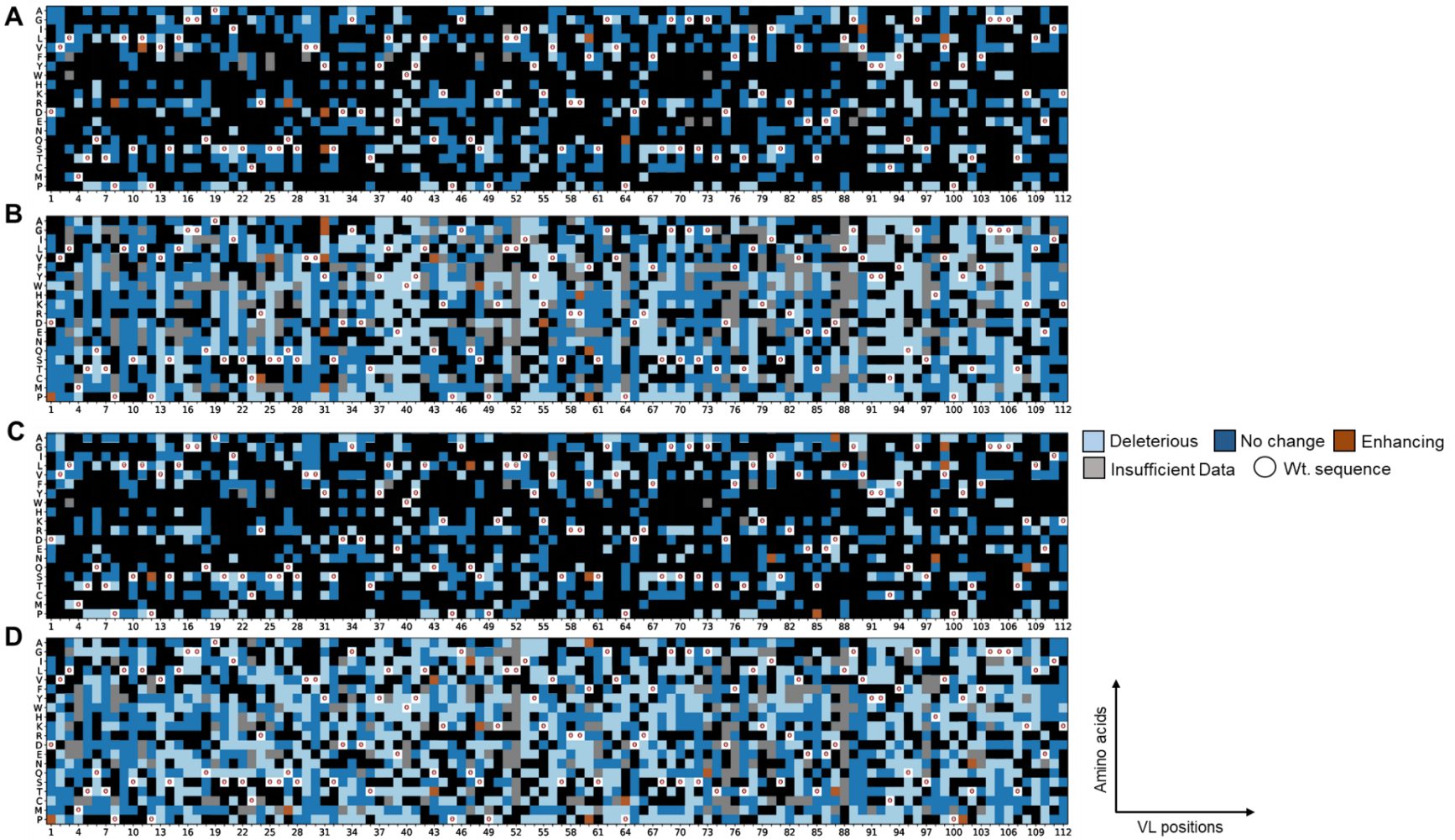

Figure S14. Heat maps highlighting functional landscapes for single vs. double/triple DNA base substitutions needed for VL single amino acid variants. Single nucleotide substitutions (with multiple nucleotide substitutions blackened-out) for VL-MNN libraries against BG505-FP8v1. **(B)** Heat map for multiple (≥2) nucleotide substitutions (single nucleotide substitutions blackened-out) for VL-MNN libraries against BG505-FP8v1. **(C)** Single nucleotide substitutions (with multiple nucleotide substitutions blackened-out) for VL-MNN libraries against BG505-FP8v2. **(D)** Heat map for multiple (≥2) nucleotide substitutions (single nucleotide substitutions blackened-out) for VL-MNN libraries against BG505-FP8v2.

Supplementary Tables

Table S1. (A) *E. coli* and **(B)** Yeast transformation experiments statistics for generating vFP16.02 VH_NNK and VL_MNN libraries.

A

E. coli Transformation								
Library Name	**Transformation**	**Theoretical Size***	**# Transformants**	**Fold coverage**	**# Colonies sequenced**	**Single mutations**	**Adjusted library size**	**Adjusted fold coverage (on DNA level)**
VL-MNN lib.	1	3,584	8.09E+05	225	19	11	4.68E+05	131
VL-MNN lib.	2	3,584	1.37E+06	382	20	9	6.17E+05	172
VH-NNK lib.	1	3,776	6.53E+06	1,729	18	9	3.27E+06	865

B

Yeast transformation			
Library	**Theoretical size***	**#Transformants**	**Fold coverage (on DNA level)**
VL-SSM lib.	3,584	2.20E+06	614
VH-SSM lib.	3,776	1.90E+06	503

*Theoretical library size on DNA level was calculated as length of VH/VL gene×No. of codons. For VH gene, the theoretical library size: 118×32=3,776; VL gene=112×32=3,584.

Table S2. vFP16.02 single mutational landscape analysis statistics.

Variable Region	Amino acid length	Theoretical size
VH	118	2,242
VL	112	2,128

Mutation Category	VH_NNK vs. BG505-FP8v1	% of mutations	VH_NNK vs. BG505-FP8v2	% of mutations	VL_MNN vs. BG505-FP8v1	% of mutations	VL_MNN vs. BG505-FP8v2	% of mutations	Adjusted Average %
Enhancing	35	1.56	18	0.803	23	1.08	23	1.08	1.30
No Change	1,400	62.4	1,222	54.5	1,066	50.1	1,014	47.7	59.6
Deleterious	617	27.5	813	36.3	756	35.5	878	41.3	39.1
No Data or Minor Effect	190	8.47	189	8.43	283	13.3	213	10.0	
Total	**2,242**	**100**	**2,242**	**100**	**2,128**	**100**	**2,128**	**100**	**100**

Statistics calculated using Round 2 sorted samples.

Table S3. Summary of single/combination mutations selected after FACS screening with their biochemical characteristics.

	Mutation	Structural location	BG505-FP8v1 High Aff. Rd3 ER	BG505-FP8v2 High Aff. Rd2 ER	Fusion Peptide			BG505.DS.SOSIP			Melting Temperature (ºC)	IC_{50} (BG505.W6.M.C2.SG3) (µg/mL)	IC_{50} (Geomean, 10 viruses) (µg/mL)	IC_{80} (Geomean, 10 viruses) (µg/mL)	Fold Change in neutralization*
					K_D (nM)	K_{on} ($M^{-1}s^{-1}$)	K_{off} (s^{-1})	K_D (nM)	K_{on} ($M^{-1}s^{-1}$)	K_{off} (s^{-1})					
Heavy chain single variants	vFP16.02		0.0116	0.717	1.07	5.79E+05	6.19E-04	0.113	2.73E+06	3.09E-04	70.4	4.31	0.99	5.38	1.00
	Q39A	FR2	0.0	1.39	0.11	6.16E+05	6.75E-05	0.048	4.60E+06	2.22E-04	69.8	0.976	0.55	2.70	1.80
	Q39S	FR2	2.07	0.146	0.098	5.41E+05	5.30E-05	0.0563	3.86E+06	2.18E-04	70.2	1.33	1.15	4.44	0.86
	Q39W	FR2	26.1	289.5	<0.001	7.10E+05	<1.0E-07	0.0595	4.13E+06	2.46E-04	71.0	2.15	0.60	3.10	1.65
	T40N	FR2	1.97	0.0	0.14	6.57E+05	8.99E-05	0.084	3.73E+06	3.14E-04	70.4	3.67	1.10	5.79	0.90
	P41G	FR2	9.31	0.003	<0.001	5.49E+05	<1.0E-07	0.0512	5.27E+06	2.70E-04	70.7	2.51	0.77	3.66	1.29
	V42G	FR2	387.0	260.2	0.073	5.65E+05	4.15E-05	0.089	3.64E+06	3.24E-04	71.5	4.96	0.99	4.94	1.00
	V42E	FR2	10.9	6.58	<0.001	3.36E+05	<1.0E-07	0.126	2.58E+06	3.25E-04	71.0	6.24	1.48	6.30	0.67
	S60V	FR3	2.08	0.0	0.16	5.42E+05	8.57E-05	0.190	4.58E+06	8.70E-04	72.3	12.4	2.33	13.2	0.42
	A92G	FR3	12.1	207.1	0.43	5.88E+05	2.51E-04	0.0587	3.44E+06	2.02E-04	72.0	1.24	0.66	3.34	1.50
	F102G	CDR3	3.66	0.00295	3.05	3.81E+05	1.16E-03	3.00	1.06E+06	3.15E-02	71.9	>50	50.0	50.0	N.N.
	V107G	CDR3	0.0	1.28	0.57	5.63E+05	3.20E-04	0.23	3.09E+06	7.15E-04	66.7	>50	4.52	19.09	0.22
Light chain single variants	Y31A	CDR1	880.3	0.111	2.15	4.19E+05	9.02E-04	0.944	7.98E+06	7.53E-03	73.5	>50	12.68	28.1	0.08
	Y31D	CDR1	6.33	0.209	2.03	1.05E+06	2.12E-03	8.45	4.83E+07	4.08E-01	69.9	>50	50.0	50.0	N.N.
	Y31S	CDR1	5.57	0.041	1.31	6.87E+05	8.99E-04	2.20	1.07E+07	2.36E-02	72.8	>50	50.0	50.0	N.N.
	Y31E	CDR1	5.12	0.029	2.47	6.38E+05	1.58E-03	3.58	4.16E+07	1.49E-01	68.1	>50	25.73	29.4	0.04
	Y31V	CDR1	4.61	0.0	1.21	4.92E+05	5.95E-04	1.26	5.05E+06	6.35E-03	69.1	>50	47.96	50.0	N.N.
	Q43T	FR2	0.004	68.4	<0.001	1.58E+05	<1.0E-07	0.065	4.86E+06	3.14E-04	68.2	1.90	0.66	3.42	1.50
	Q43V	FR2	34.5	112.3	<0.001	3.58E+05	2.33E-04	0.087	4.07E+06	3.54E-04	71.0	2.62	0.92	4.76	1.08
	S48K	FR2	0.0	97.8	<0.001	6.05E+05	<1.0E-07	0.061	4.27E+06	2.60E-04	68.7	1.14	0.36	1.43	2.75
	F60A	FR3	0.071	10.2	0.85	4.69E+05	3.97E-04	0.197	2.19E+06	4.33E-04	72.1	15.3	2.16	13.08	0.46
	F60P	FR3	53.8	153.1	<0.001	8.23E+05	<1.0E-07	0.106	3.52E+06	3.72E-04	71.1	4.16	0.95	4.59	1.04
	T85P	FR3	0.008	3.95	0.036	4.65E+05	1.69E-05	0.0869	3.12E+06	2.71E-04	71.6	4.10	1.02	5.95	0.97
	V99L	CDR3	24.2	100.4	<0.001	7.81E+05	<1.0E-07	4.20	7.92E+06	3.34E-02	70.5	2.36	0.74	3.11	1.34

Fold change calculated as IC_{50} Geomean wt/ IC_{50} Geomean mutant.
N.N.= Non Neutralizing

Neutralization potency (µg/mL)					
<0.001	.001-.01	.01-.100	.100-1.00	1.00-10.0	>10.0

Table S4. Neutralization IC_{80} of vFP16.02 single variants against a panel of 10 viral isolates.

			IC80 (µg/ml)											
		clade	A	A	A	AE	AG	B	B	BC	C	C	C	C
	Antibody	virus	BG505.W6M.C2.SG3	KER2008.12.SG3	Q23.17.SG3	CNE56.SG3	T278-50.SG3	3988.25.SG3	BL01.DG.SG3	CNE19.SG3	0077.v1.c16.SG3	25710-2.43.SG3	286.36.SG3	6644.V2.C33.SG3
Heavy Chain Single Variants	vFP16.02		15.0	19.9	34.0	3.00	>50	>50	15.8	0.732	3.87	0.724	>50	>50
	Q39A		4.02	4.15	4.74	1.26	>50	>50	4.23	0.264	1.03	0.420	8.49	>50
	Q39S		3.49	4.08	3.82	1.35	>50	>50	3.52	>50	1.27	0.414	8.67	>50
	Q39W		6.20	5.99	3.64	1.53	>50	>50	4.45	0.262	1.46	0.341	13.6	>50
	T40N		11.5	15.6	10.6	2.96	>50	>50	12.2	0.279	2.37	0.629	29.9	>50
	P41G		7.40	7.32	7.10	1.73	>50	>50	4.59	0.327	1.61	0.406	13.3	>50
	V42G		11.7	10.9	8.28	1.64	>50	>50	8.04	0.420	2.16	0.459	29.7	>50
	V42E		14.5	16.5	16.6	1.86	>50	>50	12.8	0.532	2.8	0.588	23.7	>50
	S60V		>50	>50	>50	4.32	>50	18.8	38.2	1.76	5.15	0.920	>50	>50
	A92G		5.13	5.20	5.82	1.59	>50	>50	5.34	0.318	1.48	0.483	11.4	>50
	F102G		>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50
	V107G		>50	>50	>50	9.14	>50	>50	>50	2.66	10.4	1.63	>50	>50
Light Chain Single Variants	Y31A		>50	>50	>50	>50	>50	>50	>50	0.269	34.1	42.9	>50	>50
	Y31D		>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50
	Y31E		>50	>50	>50	>50	>50	>50	>50	0.244	>50	>50	>50	>50
	Y31S		>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50
	Y31V		>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50	>50
	Q43T		6.59	5.51	7.08	1.73	>50	>50	5.9	0.266	1.11	0.428	13.1	>50
	Q43V		8.42	8.98	7.63	2.40	>50	>50	7.07	0.420	1.93	0.593	25.4	>50
	S48K		3.65	3.51	4.48	0.745	>40	>40	2.47	0.161	0.773	0.220	8.89	>40
	F60A		>50	37.9	>50	3.86	>50	>50	41.2	1.02	4.54	0.840	>50	>50
	F60P		16.1	10.7	>20	1.88	nd	>20	7.75	0.343	1.53	0.395	>20	nd
	T85P		13.2	13.7	14.9	2.65	>50	>50	12.5	0.461	1.92	0.736	19.2	>50
	V99L		5.59	9.32	5.59	1.10	>20	>20	3.83	0.347	1.19	0.413	>20	>20

Neutralization potency (µg/mL)

<0.001	.001-.01	.01-.100	.100-1.00	1.00-10.0	>10.0

Table S5. Combination mutational variants designed for further improvement of vFP16.02 by rationally combining improved single variants.

Antibody	Heavy Chain Mutations	Light Chain Mutations	IC_{50} GeoMean (10 virus panel; µg/mL)	IC_{50} (BG505.W6M.C2 .SG3; µg/mL)
vFP16.02 wt	-	-	0.950	3.57
vFP16.02 combo01	VH-Q39W	VL-V99L	0.495	1.04
vFP16.02 combo02	VH-A92G	VL-F60P	0.450	1.14
vFP16.02 combo03	VH-P41G, VH-Q39A	VL-S48K	0.321	0.50
vFP16.02 combo04	VH-P41G, VH-V42G	VL-V99L	0.526	0.68
vFP16.02 combo05	VH-Q39A, VH-P41G, VH-V42G	VL-S48K	0.514	1.16
vFP16.02 combo06	VH-P41G, VH-V42N	VL-V99L	0.567	1.37
vFP16.02 combo07	VH-P41G, VH-V42N	VL-F60P	0.587	2.00
vFP16.02 combo08	VH-Q39A	VL-F60P, VL-V99L	0.439	1.02
vFP16.02 combo09	VH-Q39A	VL-F60P, VL-V99L, VL-S48K	0.596	1.11
vFP16.02 combo10	VH-P41G, VH-V42G	VL-F60P, VL-S48K	1.01	3.64
vFP16.02 combo11	VH-Q39A	VL-S48K	0.614	1.50
vFP16.02 combo12	VH-Q39W	VL-S48K	0.763	1.31

Neutralization potency (µg/mL)

<0.001	.001-.01	.01-.100	.100-1.00	1.00-10.0	>10.0

Table S6. Neutralization IC_{50} potency of combination variants of vFP16.02 against a panel of 10 isolates.

	Clade	A	A	A	AE	AG	B	B	BC	C	C	C	C
Antibody	**Virus**	**BG505.W6M.C2.SG3**	**KER2008.12.SG3**	**Q23.17.SG3**	**CNE56.SG3**	**T278-50.SG3**	**3988.25.SG3**	**BL01.DG.SG3**	**CNE19.SG3**	**0077.v1.c16.SG3**	**25710-2.43.SG3**	**286.36.SG3**	**6644.V2.C33.SG3**
vFP16.02		3.57	0.432	0.416	0.793	>50	15.3	2.22	0.105	0.999	0.124	2.65	>50
vFP16.02 combo01		1.04	0.389	0.290	0.484	>50	5.92	0.696	0.085	0.458	0.088	1.11	>50
vFP16.02 combo02		1.14	0.127	0.210	0.450	>50	5.27	1.10	0.097	0.392	0.102	1.11	>50
vFP16.02 combo03		0.496	0.117	0.160	0.308	>50	3.02	0.692	0.087	0.359	0.064	0.963	>50
vFP16.02 combo04		0.679	0.541	0.212	0.441	>50	7.65	0.868	0.087	0.511	0.104	1.55	>50
vFP16.02 combo05		1.16	0.524	0.460	0.519	>50	4.00	0.592	0.066	0.479	0.118	1.01	>50
vFP16.02 combo06		1.37	0.382	0.387	0.369	>50	6.28	0.947	0.087	0.622	0.117	1.22	>50
vFP16.02 combo07		2.00	0.139	0.317	0.217	>50	7.16	1.97	0.121	0.519	0.103	2.79	>50
vFP16.02 combo08		1.02	0.203	0.239	0.226	>50	6.83	1.09	0.081	0.367	0.076	1.40	>50
vFP16.02 combo9		1.11	0.450	0.262	0.378	>50	10.7	1.25	0.080	0.598	0.113	1.57	>50
vFP16.02 combo10		3.64	1.68	0.693	1.12	>50	7.32	1.75	0.105	0.641	0.168	1.73	>50

Neutralization potency (μg/mL)

<0.001	.001-.01	.01-.100	.100-1.00	1.00-10.0	>10.0

Table S7. Neutralization IC_{80} potency of combination variants of vFP16.02 against a panel of 10 isolates.

	Clade	A	A	A	AE	AG	B	B	BC	C	C	C	C
Antibody	**Virus**	**BG505.W6M.C2.SG3**	**KER2008.12.SG3**	**Q23.17.SG3**	**CNE56.SG3**	**T278-50.SG3**	**3988.25.SG3**	**BL01.DG.SG3**	**CNE19.SG3**	**0077.v1.c16.SG3**	**25710-2.43.SG3**	**286.36.SG3**	**6644.V2.C33.SG3**
vFP16.02		13.6	7.56	12.3	2.91	>50	>50	8.65	0.471	2.96	0.578	26.2	>50
vFP16.02 combo01		4.04	3.60	3.90	1.75	>50	>50	2.54	0.224	1.20	0.316	9.40	>50
vFP16.02 combo02		5.62	1.95	5.58	1.89	>50	>50	3.54	0.285	1.54	0.385	12.8	>50
vFP16.02 combo03		2.79	1.80	2.94	1.45	>50	>50	2.68	0.263	1.28	0.330	6.03	>50
vFP16.02 combo04		5.00	6.44	4.42	1.59	>50	>50	2.93	0.284	1.30	0.368	11.8	>50
vFP16.02 combo05		3.85	3.86	3.13	1.69	>50	>50	2.11	0.210	1.32	0.372	6.09	>50
vFP16.02 combo06		4.55	5.00	5.79	1.57	>50	>50	3.50	0.276	1.55	0.387	10.6	>50
vFP16.02 combo07		18.1	5.87	10.6	1.60	>50	>50	7.71	0.398	1.94	0.414	32.3	>50
vFP16.02 combo08		6.97	5.04	5.00	1.40	>50	>50	4.07	0.290	1.34	0.319	19.5	>50
vFP16.02 combo9		13.5	7.03	7.48	1.93	>50	>50	4.93	0.320	1.86	0.393	24.1	>50
vFP16.02 combo10		11.9	15.6	16.1	2.10	>50	>50	6.42	0.363	1.90	0.406	31.6	>50

Neutralization potency (µg/mL)

<0.001	.001-.01	.01-.100	.100-1.00	1.00-10.0	>10.0

Table S8. Deep sequencing statistics.

Sample Name	# Raw reads	# Linked reads	# Quality fitered reads (q30p90)	Reads with complete gene without stop codons	# Sequences with ONE aa different from wt.
vFP16.02_VH_NNK-VL-positive	3,241,989	3,054,616	2,826,860	2,658,070	1,700,701
vFP16.02_VL_MNN-VL-positive	2,576,377	2,136,732	1,844,466	1,814,269	1,067,796
vFP16.02_R1_1_VH_NNK (BG505-FP8v1-Low)	673,446	645,198	591,591	557,338	385,303
vFP16.02_R1_2_VH_NNK (BG505-FP8v1_Med)	348,053	338,366	307,947	283,781	177,595
vFP16.02_R1_3_VH_NNK (BG505-FP8v1_High)	**517,617**	**498,426**	**454,893**	**429,852**	**265,797**
vFP16.02_R1_4_VH_NNK (BG505-FP8v2_Low)	567,891	532,055	491,135	467,539	331,217
vFP16.02_R1_5_VH_NNK (BG505-FP8v2_Med)	590,894	557,808	514,843	481,540	299,570
vFP16.02_R1_6_VH_NNK (BG505-FP8v2_High)	**804,343**	**757,036**	**699,032**	**669,455**	**419,885**
vFP16.02_R1_7_VL_MNN (BG505-FP8v1_Low)	488,412	433,978	373,132	360,458	257,481
vFP16.02_R1_8_VL_MNN (BG505-FP8v1_Med)	884,676	693,948	572,246	555,783	256,944
vFP16.02_R1_9_VL_MNN (BG505-FP8v1_High)	**407,564**	**380,974**	**324,721**	**312,193**	**169,963**
vFP16.02_R1_10_VL_MNN (BG505-FP8v2_Low)	581,428	414,670	343,176	331,988	232,360
vFP16.02_R1_11_VL_MNN (BG505-FP8v2_Med)	591,175	526,837	419,441	406,551	223,437
vFP16.02_R1_12_VL_MNN (BG505-FP8v2_High)	**499,312**	**453,617**	**392,878**	**379,549**	**202,867**
vFP16.02_R2_1_VH_NNK (BG505-FP8v1_Low)	866,906	796,257	732,397	717,493	511,593
vFP16.02_R2_2_VH_NNK (BG505-FP8v1_Med)	328,955	324,790	301,227	292,011	166,036
vFP16.02_R2_3_VH_NNK (BG505-FP8v1_High)	**511,136**	**463,809**	**428,907**	**422,333**	**332,222**
vFP16.02_R2_4_VH_NNK (BG505-FP8v2_Low)	347,117	315,655	287,026	281,398	205,843
vFP16.02_R2_5_VH_NNK (BG505-FP8v2_Med)	592,231	575,739	521,794	502,826	295,444
vFP16.02_R2_6_VH_NNK (BG505-FP8v2_High)	**82,348**	**78,029**	**65,785**	**64,479**	**49,347**
vFP16.02_R2_7_VL_MNN (BG505-FP8v1_Low)	724,647	628,030	528,822	520,943	413,753
vFP16.02_R2_8_VL_MNN (BG505-FP8v1_Med)	464,077	370,679	319,382	315,124	111,016
vFP16.02_R2_9_VL_MNN (BG505-FP8v1_High)	**299,741**	**286,151**	**238,854**	**234,531**	**193,671**
vFP16.02_R2_10_VL_MNN (BG505-FP8v2_Low)	233,219	208,417	179,407	176,239	139,543
vFP16.02_R2_11_VL_MNN (BG505-FP8v2_Med)	560,550	512,003	433,992	426,913	206,966
vFP16.02_R2_12_VL_MNN (BG505-FP8v2_High)	**761,092**	**624,515**	**522,391**	**514,799**	**318,880**
vFP16.02_R3_1_VH_NNK (BG505-FP8v1-Low)	528,178	430,702	367,392	359,883	260,093
vFP16.02_R3_2_VH_NNK (BG505-FP8v1_Med)	881,646	738,056	633,377	620,559	325,397
vFP16.02_R3_3_VH_NNK (BG505-FP8v1_High)	**498,302**	**426,056**	**357,975**	**352,294**	**277,510**
vFP16.02_R3_4_VH_NNK (BG505-FP8v2_Low)	717,190	579,841	499,838	491,717	363,749
vFP16.02_R3_5_VH_NNK (BG505-FP8v2_Med)	458,494	394,648	339,124	331,619	182,162
vFP16.02_R3_6_VH_NNK (BG505-FP8v2_High)	**711,398**	**600,709**	**518,094**	**510,987**	**433,210**
vFP16.02_R3_7_VL_MNN (BG505-FP8v1_Low)	761,105	552,164	452,682	444,620	353,228
vFP16.02_R3_8_VL_MNN (BG505-FP8v1_Med)	1,022,851	751,986	622,681	613,956	195,251
vFP16.02_R3_9_VL_MNN (BG505-FP8v1_High)	**606,295**	**532,021**	**400,546**	**392,892**	**320,500**
vFP16.02_R3_10_VL_MNN (BG505-FP8v2_Low)	720,427	541,323	451,844	442,954	345,156
vFP16.02_R3_11_VL_MNN (BG505-FP8v2_Med)	522,460	398,530	305,657	300,040	126,863
vFP16.02_R3_12_VL_MNN (BG505-FP8v2_High)	**777,876**	**499,718**	**410,887**	**403,487**	**61,306**

Table S9. Crystal structure data collection and refinement statistics.

	VL-S48K	VL-F60P
PDB access code	6WWC	6WX2
Wavelength (Å)	1	1
Resolution range	34.8 - 2.6 (2.7- 2.6)	35.1 - 2.4 (2.5 - 2.4)
Space group	P 2_1	P 2_1 2_1 2_1
Unit cell	68.6 70.5 93.3 90 94.8 90	73.2 111.7 123.8 90 90 90
Unique reflections	23878 (606)	39868 (3090)
Multiplicity	3.1 (2.9)	11.0 (9.5)
Completeness (%)	83.1 (21.1)	97.0 (76.4)
Mean I/sigma(I)	9.2 (2.0)	25.4 (2.0)
Wilson B-factor	39.93	47.27
R-merge	0.124 (0.438)	0.155 (0.806)
R-meas	0.150 (0.527)	0.162 (0.852)
R-pim	0.083 (0.289)	0.048 (0.269)
CC1/2	0.988 (0.857)	0.999 (0.913)
Reflections used in refinement	23874 (605)	39733 (3073)
Reflections used for R-free	1108 (26)	1992 (154)
R-work	0.21 (0.31)	0.20 (0.27)
R-free	0.27 (0.49)	0.24 (0.32)
Number of non-hydrogen atoms	6814	6783
macromolecules	6706	6645
solvent	108	138
Protein residues	878	872
RMS(bonds)	0.006	0.009
RMS(angles)	1.32	1.15
Ramachandran favored (%)	96.98	98.83
Ramachandran allowed (%)	2.78	1.17
Ramachandran outliers (%)	0.23	0
Rotamer outliers (%)	0.13	7.76
Clashscore	9.44	6.4
Average B-factor	42.98	51.45
macromolecules	43.05	51.5
solvent	38.8	48.69

Statistics for the highest-resolution shell are shown in parentheses.

SI References

1. B. Wang, *et al.*, Functional interrogation and mining of natively paired human V H :V L antibody repertoires. *Nat Biotechnol* **36**, 152–155 (2018).

2. E. E. Wrenbeck, *et al.*, Plasmid-based one-pot saturation mutagenesis. *Nat Methods* **13**, 928–930 (2016).

3. A. E. Wentz, E. V. Shusta, A Novel High-Throughput Screen Reveals Yeast Genes That Increase Secretion of Heterologous Proteins. *Appl Environ Microbiol* **73**, 1189–1198 (2007).

4. L. Benatuil, J. M. Perez, J. Belk, C.-M. Hsieh, An improved yeast transformation method for the generation of very large human antibody libraries. *Protein Eng Des Sel* **23**, 155–159 (2010).

5. J. D. Hunter, Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering* **9**, 90–95 (2007).

6. A. Kassambara, ggpubr: “ggplot2” Based Publication Ready Plots (2020).

7. A. Cui, *et al.*, A Model of Somatic Hypermutation Targeting in Mice Based on High-Throughput Ig Sequencing Data. *J. Immunol.* **197**, 3566–3574 (2016).

8. Z. Sheng, *et al.*, Gene-Specific Substitution Profiles Describe the Types and Frequencies of Amino Acid Changes during Antibody Somatic Hypermutation. *Front Immunol* **8**, 537 (2017).

9. Z. Otwinowski, W. Minor, [20] Processing of X-ray diffraction data collected in oscillation mode. *Meth. Enzymol.* **276**, 307–326 (1997).

10. A. J. McCoy, *et al.*, Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674 (2007).

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

12. P. D. Adams, *et al.*, Recent developments in the PHENIX software for automated crystallographic structure determination. *J Synchrotron Radiat* **11**, 53–55 (2004).

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

14. J. C. Phillips, *et al.*, Scalable molecular dynamics with NAMD. *J Comput Chem* **26**, 1781–1802 (2005).

15. R. B. Best, *et al.*, Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain χ(1) and χ(2) dihedral angles. *J Chem Theory Comput* **8**, 3257–3273 (2012).

16. O. Guvench, *et al.*, CHARMM additive all-atom force field for carbohydrate derivatives and its utility in polysaccharide and carbohydrate-protein modeling. *J Chem Theory Comput* **7**, 3162–3180 (2011).

17. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, Comparison of simple potential functions for simulating liquid water. *jcp* **79**, 926–935 (1983).

18. J. Dunbar, A. Fuchs, J. Shi, C. M. Deane, ABangle: characterising the VH-VL orientation in antibodies. *Protein Eng. Des. Sel.* **26**, 611–620 (2013).

19. R. Kong, *et al.*, Fusion peptide of HIV-1 as a site of vulnerability to neutralizing antibody. *Science* **352**, 828–833 (2016).