Supporting Information

Liu *et al.* 10.1073/pnas.0809543105

SI Text

Library Construction

Creation of the pSEX-GermNNK Library. A heavy-chain gene fragment containing V_H 3–23 was amplified from a synthetic Fab antibody gene (see *Materials and Methods*) with primers

VH3–23-F: 5′-TCTCGAAATCCATGGCTGAGGTG-CAGCTGTTGGAGTCTGG-3′ and

VH3–23-R: 5′-TCTTTCGCACAGTAATATACGG-3′.

To this fragment was added a randomized CDR3 loop using overlap PCR with primers

NNK1: 5′-CCGTATATTACTGTGCGAAAGANNNKNN-KNNKNNKNNKNNKNACTACTTTGACTACTGGGG-3′ and

NNK2: 5′- AGCCATCGCGGCCGCGCTAGCTGAG-GAGACGGTGACCAGGGTTCCTTGGC-CCCAGTAGTCAAAG-3′,

where N = A, T, G, or C and K = G or T. The final gene product, which contains the full heavy-chain library, was then amplified with primers

HC1: 5′-CGGCCATGGCTGAGGTGCAGCTGTTG-GAGTCTGG-3′ and

HC2: 5′-CTTCAAGCTTTGGGGCGGATGCACTC-CCTGAGGAGACGGTGACCAGGGTTCCTTGG-3′.

The product was digested at restriction sites NcoI and HindIII and inserted into a similarly digested pSEX-GermTAT to create the library pSEX-GermNNK. The ligation mixture was precipitated and the product was transformed into electrocompetent Top10 F' cells to afford 5×10^8 total transformants. After overnight growth in 2YT supplemented with 100-μg/ml ampicillin, 30-μg/ml tetracycline, and 1% glucose, the supercoiled DNA was isolated. This DNA was then transformed into Keto-X-*E. coli*, Bpy-X-*E. coli*, SY-X-*E. coli*, and Boro-X-*E. coli* for phage-display studies on the population.

Creation of a 412d Library. To create a 412d library with residues Pro-101, Tys104, Asn-105, Tys107, Ala-108, Pro-109, Gly-112, and Met-113 randomized, we first amplified a heavy-chain fragment from pSEX-412d2TAG with the primers

412dlib1: 5′-CCGCTGGCTTGCTGCTGCTG-3′ and

412dlib2: 5′-GTAAGGGCTCGCACAGTAAAATACG-GCC-3′.

The resulting product was then extended with overlap PCR using the primers

412dlib3: 5′-GCCGTATTTTACTGTGCGAGCCCTTAC-NNKAATGACNNKNNKGACNNKNNKNNK-GAGGAGNNKNNKAGCTGGTACTTCGATCTCTG-3′ and

412dlib4: CAGAGATCGAAGTACCAGCT,

where N = A, T, G, or C and K = G or T. A second fragment was amplified from pSEX-412d2TAG using the primers

412dlib5: AGCTGGTACTTCGATCTCTG and

412dlib6: CTCTGATATCTTTGGATCCA

and the two fragments were assembled by overlap PCR to give the 412d gene library. This library was then digested with NcoI and BamHI and inserted into a similarly digested pSEX81 vector. The ligation product was precipitated with tRNA assistance and transformed into Top10 F' cells, affording 2×10^8 total transformants. After overnight growth in 2YT supplemented with 100-μg/ml ampicillin, 30-μg/ml tetracycline, and 1% glucose, the supercoiled DNA was isolated. This DNA was then transformed into SY-X-*E. coli* for phage display and selection of anti-gp120 antibodies.

Creation of a Naïve Germline Antibody Library for Selection of Anti-gp120 Antibodies. A mixture of germline V_H fragments was amplified from human cDNA using primers

VH-Mix-F: 5′-TCTCGAAATCCATGGCTCAGGTG-CAGCTGGTGCAGTCTGG-3′ and

VH-Mix-R: 5′-TCTCTCGCACAGTAATACACGGCCG-3′.

Annealing temperature for this PCR was 56°C. To these fragments were added a randomized CDR3 loop using overlap PCR with primers:

NNK1: 5′-CGGCCGTGTATTACTGTGCGAGAGANNN-KNNKNNKNNKNNKNNKNACTACTTTGACTACTGG-GG-3′ and

NNK2: 5′-AGCCATCGCGGCCGCGCTAGCTGAG-GAGACGGTGACCAGGGTTCCTTGGCCCCAGTAGT-CAAAG-3′,

where N = A, T, G, or C and K = G or T. The final gene products were then amplified with primers

VH66CC-F: 5′-CCGGCCATGGCTCAGGTGCAGCTGGT-GCAGTCTGG-3′ and

VH66CC-R: 5′-CTTCAAGCTTTGGGGCGGATGCACT-CCCTGAGGAGACGGTGACCAGGGTTCCT-3′.

This PCR library was digested with NcoI and HindIII and inserted into a similarly digested pSEX-GermTAT to create a library of naïve germline scFvs. The ligation mixture was precipitated and the product was transformed into electrocompetent Top10 F' cells to obtain 2×10^9 total transformants. After overnight growth in 2YT supplemented with 100-μg/ml ampicillin, 30-μg/ml tetracycline, and 1% glucose, the supercoiled DNA was isolated. This DNA was then transformed into SY-X-*E. coli* for phage display and selection of anti-gp120 antibodies.

Gp120-Binding Selection Using a Doped Library. The general gp120-binding selection procedure was used (see *Materials and Methods*). Specifically for this selection, the following washing protocols were used:

Round 1: Washed 10 times with/200-μl PBST per wash; ≈1 min per wash.

Round 2: Washed 12 times with 200-μl PBST per wash; ≈1 min per wash.

Round 3: Washed 14 times with 200-μl PBST per wash; ≈1 min per wash.

Round 4: Washed 15 times with 200-μl PBST per wash; ≈1 min per wash.

For this selection, the amounts of loaded and eluted phage throughout the rounds are listed in Table S5. Also listed is the amount of phage present in the final wash before elution, if measured. For this selection, the percent of phage containing at least 1 TAG and the percent of phage containing the 412d2TAG doped clone are listed in Table S6.

Gp120-Binding Selection Using a Naïve Germline Library. The general gp120-binding selection procedure was used (see *Materials and Methods*). Specifically for this selection, the following washing protocols were used:

Round 1: Washed 2 times with 200-μl PBST per wash; ≈1 min per wash.

Round 2: Washed 10 times with 200-μl PBST per wash and 1 time with 200-μl PBS; ≈1 min per wash.

Round 3: Washed 10 times with 200-μl PBST per wash and 1 time with 200-μl PBS; ≈1 min per wash.

Round 4: Washed 10 times with 200-μl PBST per wash and 1 time with 200-μl PBS; ≈1 min per wash.

For this selection, the amounts of loaded and eluted phage throughout the rounds are listed in Table S7. Note that the stringency of round 1 was low and, thus, the amount of eluted phage was high. This ensures that no arbitrary loss of functional clones occurs when very few copies of each clone are present in the library.

Expression of Free scFv Proteins. To express scFvs as free proteins without fusion to phage, the coding regions for scFvs 66CC8, 412d-2SY, and 412d-Y were inserted into the pBAD expression vector containing a gIII periplasmic signal sequence, and a C-terminal 6X-histidine tag using standard methods. This yielded pBAD-66CC8, pBAD-412d-2SY, and pBAD-412d-Y.

To express 412d-Y as an scFv, Top10 F' cells containing pBAD-412d-Y were grown at 37°C at 250 rpm in 2YT supplemented with 100-μg/ml ampicillin until the optical density reached 0.6, at which point scFv production was induced with 0.1% L-arabinose. The culture was then allowed to shake at 250 rpm at room temperature for 30 h, after which cells were pelleted and lysed in 8M urea because no soluble scFv was produced. Refolding using the Protein Refolding Kit (Novagen) yielded soluble scFv.

To express 412d-2SY as an scFv, Top10 F' cells containing pBAD-412d-SY along with pSUPAR6-L3–3SY, an optimized plasmid adapted for expression with unnatural amino acids that, in this case, contains the sulfotyrosine-specific synthetase and the corresponding orthogonal tRNA [Cellitti S, *et al.* (2008) *J Am Chem Soc* 130:9268–9281], were grown at 37°C at 250 rpm in 2YT supplemented with 100-μg/ml ampicillin, 30-μg/ml chloramphenicol, and 10-mM sulfotyrosine (Bachem). When the optical density reached 0.6, both synthetase and scFv production were induced with 0.2% L-arabinose. The culture was then allowed to shake at 250 rpm at room temperature for 30 h, after which cells were pelleted and lysed in 8M urea. Refolding using the Protein Refolding Kit (Novagen) yielded soluble scFv.

To express 66CC8-SY as an scFv, Top10 F' cells containing pBAD-66CC8 along with pSUPAR6-L3–3SY were grown at 37°C at 250 rpm in 2YT supplemented with 100-μg/ml ampicillin, 30-μg/ml chloramphenicol, and 10-mM sulfotyrosine (Bachem). When the optical density reached 0.6, both synthetase and scFv production were induced with 0.2% L-arabinose. The culture was then allowed to shake at 250 rpm at room temperature for 30 h, after which cells were pelleted and lysed in 8M urea. Refolding was unsuccessful as no soluble protein was recovered.

To express 66CC8-Y as an scFv, Top10 F' cells containing pBAD-66CC8 along with pCDF-JYRS, a plasmid that encodes tyrosine in response to TAG, were grown at 37°C at 250 rpm in 2YT supplemented with 100-μg/ml ampicillin and 30-μg/ml chloramphenicol. When the optical density reached 0.6, both synthetase and scFv production were induced with 0.2% L-arabinose. The culture was then allowed to shake at 250 rpm at room temperature for 30 h, after which cells were pelleted and lysed in 8M urea. Refolding was unsuccessful as no soluble protein was recovered.

Protein yield was determined by Western blot analysis using an anti-6XHis antibody (Sigma) and a standard 6X-histidine tagged protein of known concentration.

Expression and Purification of Free Fab Proteins. To convert phage-displayed scFvs into Fab format, the light chain and heavy chain variable regions for scFvs 66CC14, 66CC8, 412d-2SY, and 412d-Y were separately inserted into the pBC expression vector containing human heavy and light-chain constant regions (synthesized by Blue Heron) using standard methods. This yielded pBC-66CC14Fab, pBC-66CC8Fab, pBC-412d-2SYFab, and pBC-412d-YFab for Fab expression from a bicistronic construct under the lac promoter.

To express 66CC14 and 412d-Y as Fabs, Top10 F' cells containing either pBC-66CC14Fab or pBC-412d-YFab were grown at 37°C at 250 rpm in 2YT supplemented with 100-μg/ml ampicillin until the optical density reached 0.6, at which point Fab production was induced with 1 mM IPTG. The culture was then allowed to shake at 250 rpm at room temperature for 30 h, after which cells were pelleted and lysed three times with 1/20 culture volume of periplasmic lysis buffer (20% sucrose, 30-mM Tris-HCl, 1-mM EDTA, 1-mg/ml lysozyme, pH 7.4). The periplasmic lysate was collected for purification on protein G resin.

To express 412d-2SY and 66CC8-SY as Fabs, Top10 F' cells containing pBC-412d-SYFab or pBC-66CC8-SYFab along with pSUPAR6-L3–3SY were grown at 37°C at 250 rpm in 2YT supplemented with 100-μg/ml ampicillin, 30-μg/ml chloramphenicol, and 10-mM sulfotyrosine (Bachem). When the optical density reached 0.3, synthetase production was induced with 0.2% L-arabinose. When the optical density reached 0.6, Fab production was induced with 1-mM IPTG. The culture was then allowed to shake at 250 rpm at room temperature for 30 h, after which cells were pelleted and lysed three times with 1/20 culture volume of periplasmic lysis buffer. The periplasmic lysate was collected for purification on protein G resin.

To express 66CC8-Y as a Fab, Top10 F' cells containing pBC-66CC8Fab along with pCDF-JYRS were grown at 37°C at 250 rpm in 2YT supplemented with 100-μg/ml ampicillin and 30-μg/ml chloramphenicol. When the optical density reached 0.3, synthetase production was induced with 0.2% L-arabinose. When the optical density reached 0.6, Fab production was induced with 1 mM isopropyl-beta-D-thiogalactopyranoside (IPTG). The culture was then allowed to shake at 250 rpm at room temperature for 30 h, after which cells were pelleted and lysed three times with 1/20 culture volume of periplasmic lysis buffer. The periplasmic lysate was collected for purification on protein G resin.

To purify periplasmic lysates using protein G, 1 ml of protein G resin (Pierce) was packed into a 1-ml polypropylene column (Qiagen). After equilibration of the column with 5-ml binding buffer (50-mM Mes, 100-mM NaCl, pH 5.5), periplasmic lysate was loaded onto the column and allowed to pass through the resin by gravity flow. The column was then washed with 15-ml binding buffer and then eluted with 5-ml elution buffer (100-mM glycine, pH 2.8), which was immediately neutralized to pH 7.4. The eluted Fab was then dialyzed into PBS and concentrated for further use. Fab yield was determined by UV absorbance at $\lambda = 280$.

Determination of gp120-Binding by Fabs Using ELISA. Per sample, 0.33 μg of soluble ADA gp120 was coated onto the surface of a MaxiSorp (Nunc) microtiter plate well in 100-μl PBS for 12 h at 37°C. After blocking for 2 h with 200 μl 2% milk (Bio-Rad) in PBS, Fab proteins, previously purified and quantified according to the procedures described above, were loaded in 100 μl 2% milk PBST and incubated at 37°C for 2 h. After washing five times with PBST, an antihuman kappa light-chain antibody (Sigma) was added in 110 μl of 2% milk in PBST and incubated at 37°C for 2 h. After washing eight times with PBST, QuantaBlu fluorogenic substrate (Pierce) was added and the ELISA signal was determined using a fluorescence plate reader (SpectraMax Gemini).

Synthesis of Bipyridyl-Alanine

5-Methyl-2,2′-Bipyridine. A mixture of 2-bromo-5-methylpyridine (5.0 g, 29 mmol), 2-tributylstannylpyridine (10 g, 27 mmol), and $Pd(PPh_3)_4$ (2.0 g, 1.7 mmol) in dry toluene (250 ml) was stirred for 48 h at 110°C. The reaction mixture was filtered over Celite and evaporated under reduced pressure. The residue was dissolved in EtOAc and washed with saturated aqueous $NaHCO_3$

solution, and the organic layer was dried over $MgSO_4$ and concentrated under reduced pressure. Flash column chromatography (0.5% MeOH in CH_2Cl_2) afforded methyl-2,2′-bipyridine (3.5 g, 75%) as a colorless oil. ^{1}H-NMR (500 MHz, CDCl3): δ 2.47 (s, 3H), 7.35 (dd, 1H), 7.70 (dd, 1H), 7.87 (m, 1H), 8.36 (d, 1H), 8.43 (d, 1H), 8.58 (d, 1H), 8.74 (dd, 1H).

Diethyl 2-(2,2′-bipyridin-5-ylmethyl)-2-Acetamidomalonate. Methyl-2,2′-bipyridine (1.7 g, 10 mmol) in a mixed solvent (100 ml) of water and benzene (1:1) was irradiated and refluxed using a halogen lamp (150 W). Bromine (1.6 g, 10 mmol) was added and the mixture was heated to reflux for 30 min. The solution was concentrated under reduced pressure and dissolved in EtOAc. The organic layer was washed with saturated aqueous $NaHCO_3$ solution, dried over $MgSO_4$, and concentrated under reduced pressure to afford the crude brominated product. A mixture of diethyl acetamidomalonate (2.6 g, 12 mmol) and sodium hydride (0.48 g, 12 mmol, 60% in mineral oil) in dry DMF (50 ml) was stirred for 30 min at 0°C. To the solution was added the crude product in dry DMF (20 ml) at 0°C and the mixture was stirred for 1 h at room temperature. The reaction mixture was diluted with EtOAc (200 ml) and washed with a 10% aqueous sodium thiosulfate solution (2 × 150 ml). The organic layer was dried over $MgSO_4$ and concentrated under reduced pressure. The crude product was purified by flash column chromatography (0.5% MeOH in CH_2Cl_2) to give diethyl 2-(2,2′-bipyridin-5-ylmethyl)-2-acetamidomalonate (2.5 g, 65%) as a white solid. ^{1}H-NMR (500 MHz, CDCl3): δ 1.38 (m, 6H), 2.15 (s, 3H), 3.81 (s, 2H), 4.35 (m, 4H), 6.68 (s, 1H), 7.36 (dd, 1H), 7.55 (dd, 1H), 7.87 (m, 1H), 8.36 (d, 1H), 8.41 (d, 1H), 8.42 (d, 1H), 8.74 (dd, 1H).

Bipyridyl-Alanine: 3-(2,2′-Bipyridin-5-yl)-2-Aminopropanoic Acid. Diethyl 2-(2,2′-bipyridin-5-ylmethyl)-2-acetamidomalonate (2.5 g, 6.5 mmol) in 12 M HCl was heated to reflux for 6 h. The reaction mixture was concentrated under reduced pressure to give the product (2.3 g, 99%) as a white solid in HCl salt form. ^{1}H-NMR (500 MHz, $CDCl_3$: δ3.59 (m, 2H), 4.58 (t, 1H), 8.11 (m, 1H), 8.40 (dd, 1H), 8.67 (d, 1H), 8.70 (m, 1H), 8.83 (d, 1H), 8.99 (dd, 1H). LC-MS (ESI) for calculated $C_{13}H_{13}N_3O_2$ (M + 1) 243.1, observed 243.1.

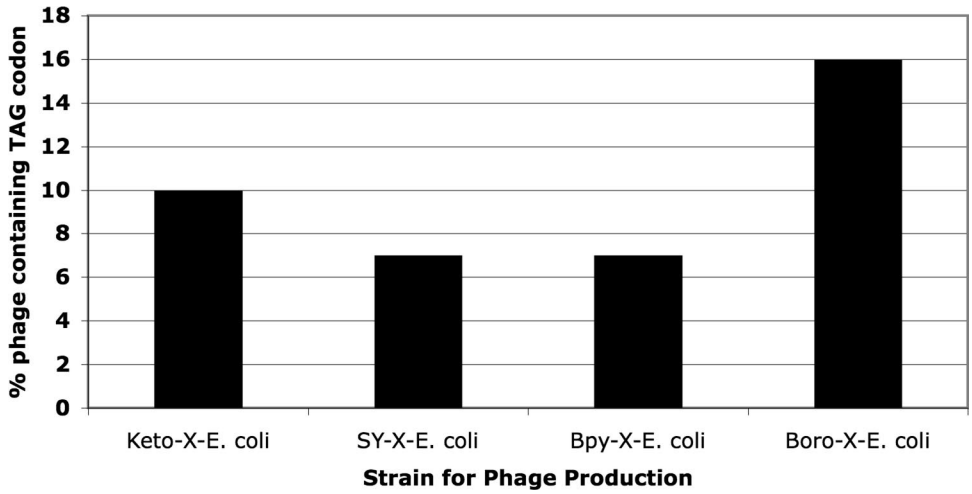

Fig. S1. Percent-phage clones containing one or more TAG codons after phage expression from the pSEX-GermNNK library in Keto-X-*E. coli*, SY-X-*E. coli*, Bpy-X-*E. coli*, or Boro-X-*E. coli* ($n = 100$). Expected value is 17.3%; deviation represents a bias in favor of sequences containing only the 20 canonical amino acids. Phage were produced under optimized conditions.

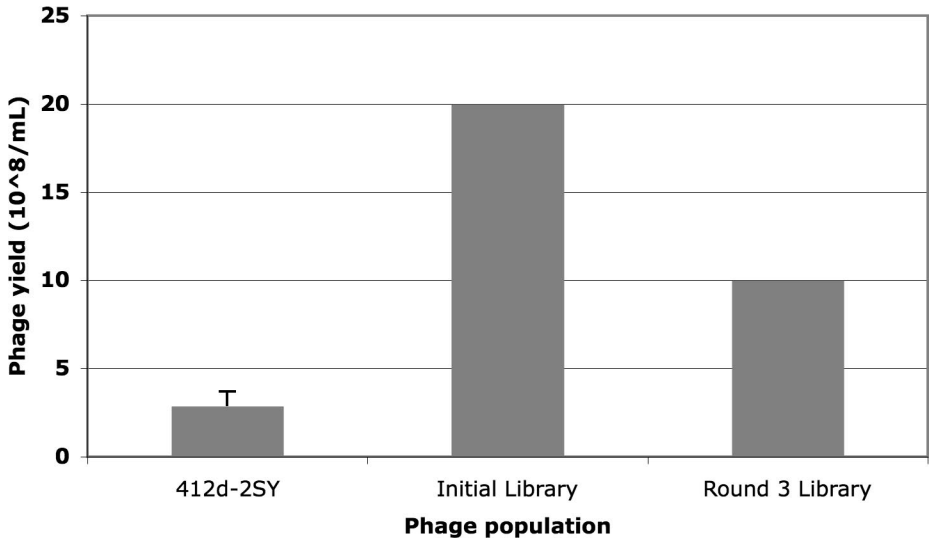

Fig. S2. Yield of phage per mL culture of 412d-2SY in comparison with phage yield from initial phage library and library at round three. All phage were produced using SY-X-*E. coli*. For 412d-2SY, titers from three separate phage preparations were averaged and error bar represents plus standard deviation. For 412d-2SY, when sulfotyrosine was omitted from the media, phage yield was $\approx 1 \times 10^6$ per mL.

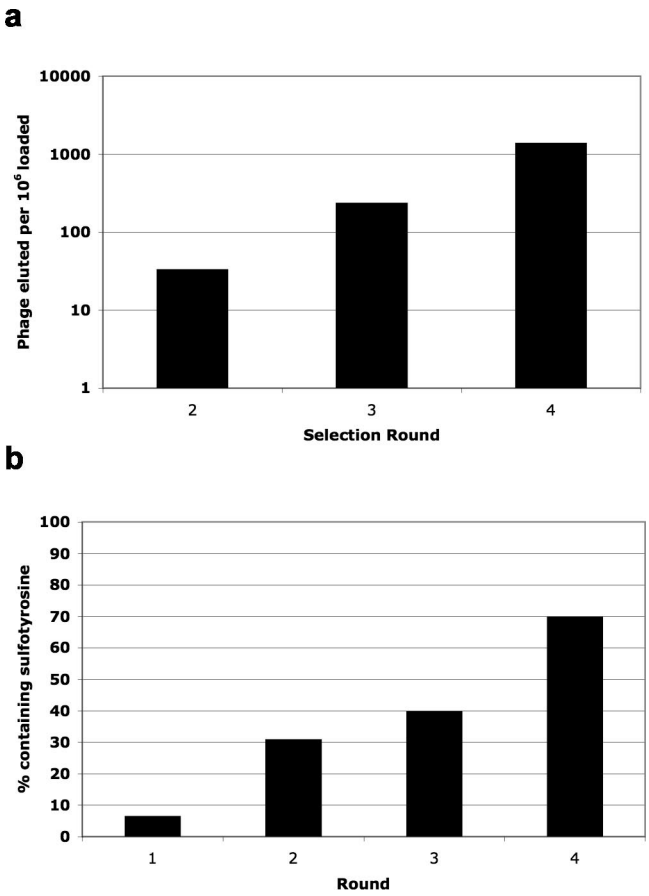

Fig. S3. (*a*) Enrichment for gp120 binding as judged by the eluted phage amount after each round. Post-round one elution titers are not shown as round one selection was done with much lower stringency than with subsequent rounds (see *SI Text*) to minimize arbitrary loss of potential hits when few copies of each clone were present. (*b*) Increase in the percent-clones containing sulfotyrosine after each round as determined by sequencing (n = 15–30).

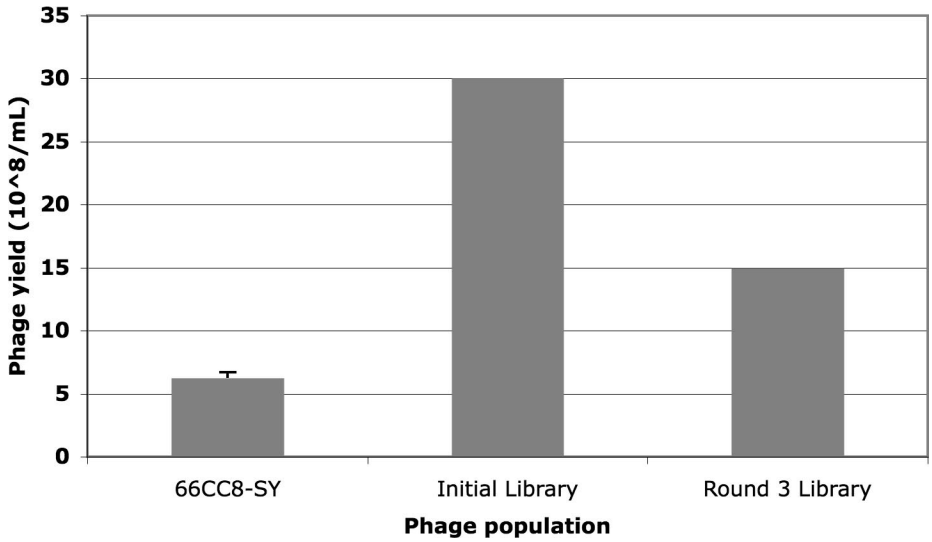

Fig. S4. Yield of phage per mL culture of 66CC8-SY in comparison with phage yield from initial phage library and library at round three. All phage were produced in SY-X-*E. coli*. For 66CC8-SY, titers from three separate phage preparations were averaged and error bar represents plus standard deviation. For 66CC8-SY, when unnatural amino acid was omitted from the media, phage yield was $\approx 5 \times 10^6$.

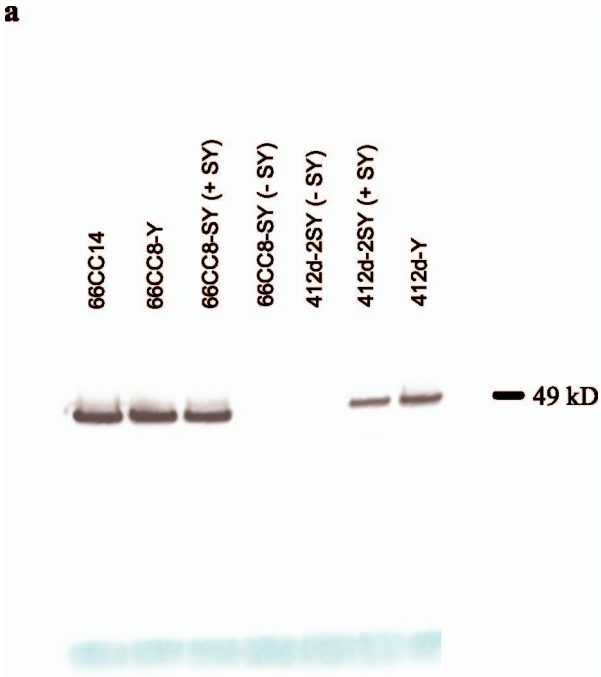

Fig. S5. (*a*) Western blot analysis of protein-G purified Fabs using antihuman kappa light chain HRP antibody developed with metal-enhanced DAB kit (Pierce). Samples were run on a denaturing PAGE gel (Invitrogen NuPAGE 4–12% Bis-Tris). For 66CC8- SY and 412d-SY, lanes corresponding to expression in the absence of sulfotyrosine are also presented to show dependence of sulfated antibody expression on the presence of sulfotyrosine. (*b*) LCMS (ESI-positive) spectra of Fabs 66CC14, 66CC8-SY, 66CC8-Y, 412d-SY, and 412d-Y. (*c*) ELISA measuring binding of gp120 by purified Fabs 412d-2SY, 412d-Y, 66CC8-SY, 66CC8, and 66CC14.

b

66CC14 ([M+H] calculated: 47698; [M+H] observed: 47697)

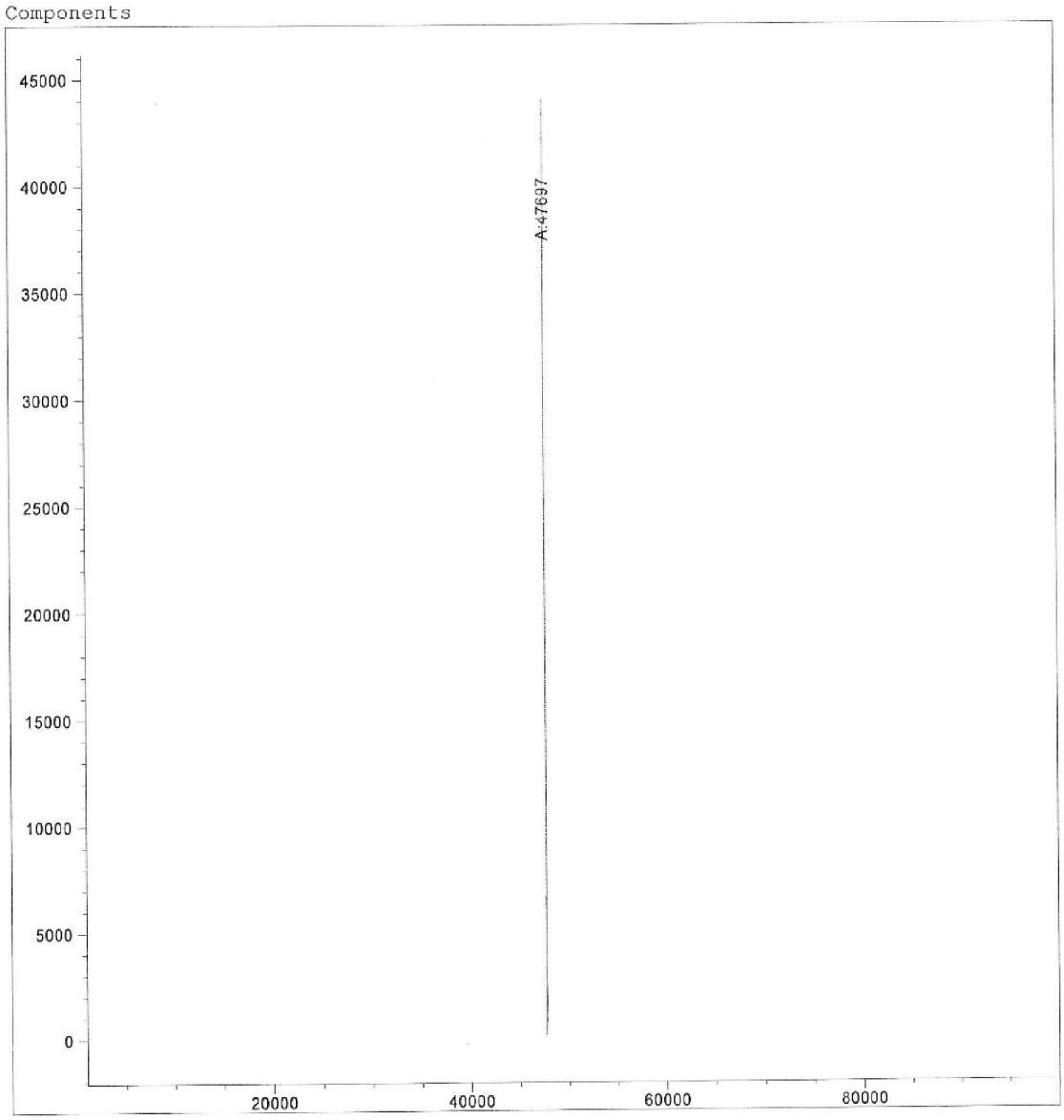

Fig. S5 (continued).

66CC8-SY ([M+H] calculated: 47626; [M+H] observed: 47621)

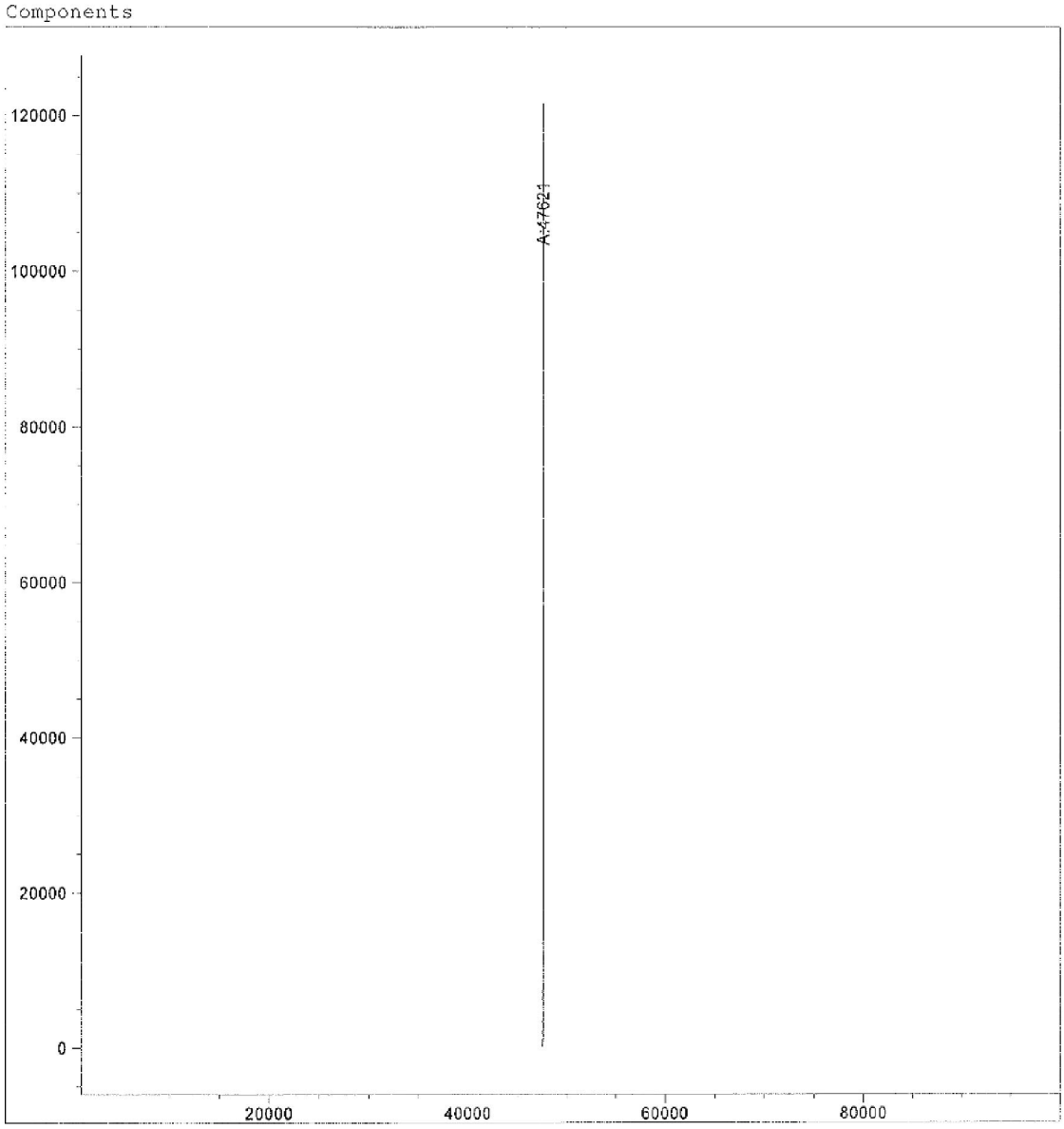

Fig. S5 (continued).

66CC8-Y ([M+H] calculated: 47546; [M+H] observed: 47540)

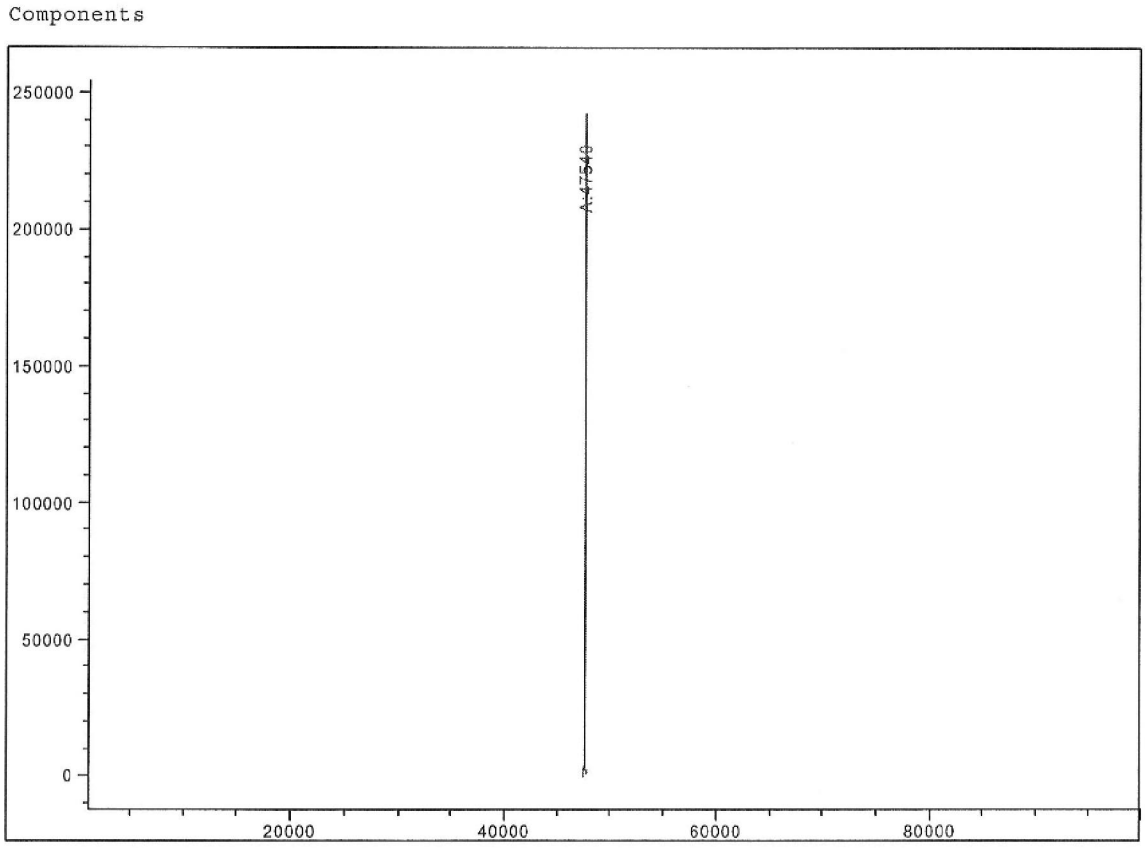

Fig. S5 (continued).

412d-2SY ([M+H] calculated: 48822; [M+H] observed: 48823)

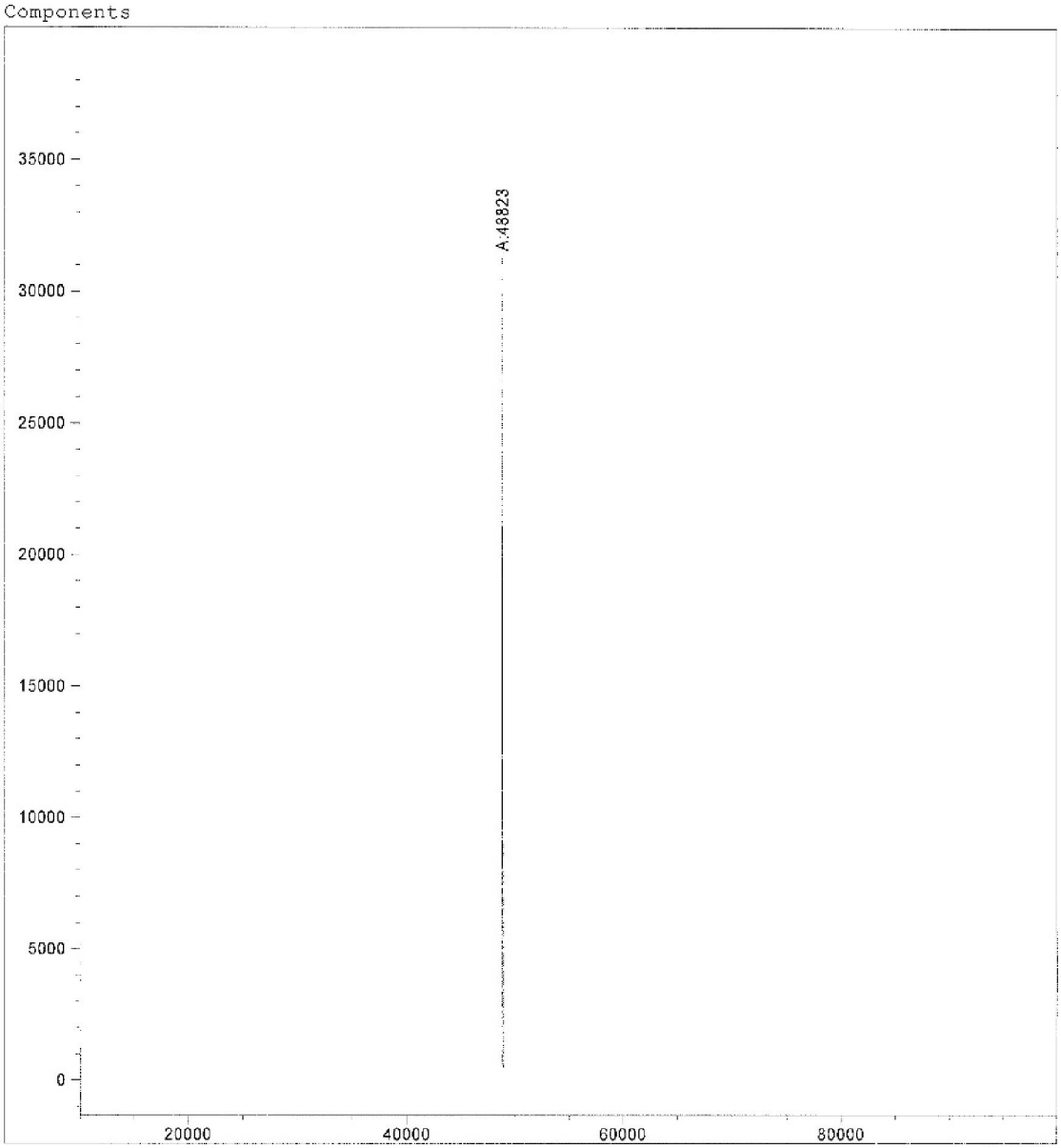

Fig. S5 (continued).

412d-Y ([M+H] calculated: 48662; [M+H] observed: 48665)

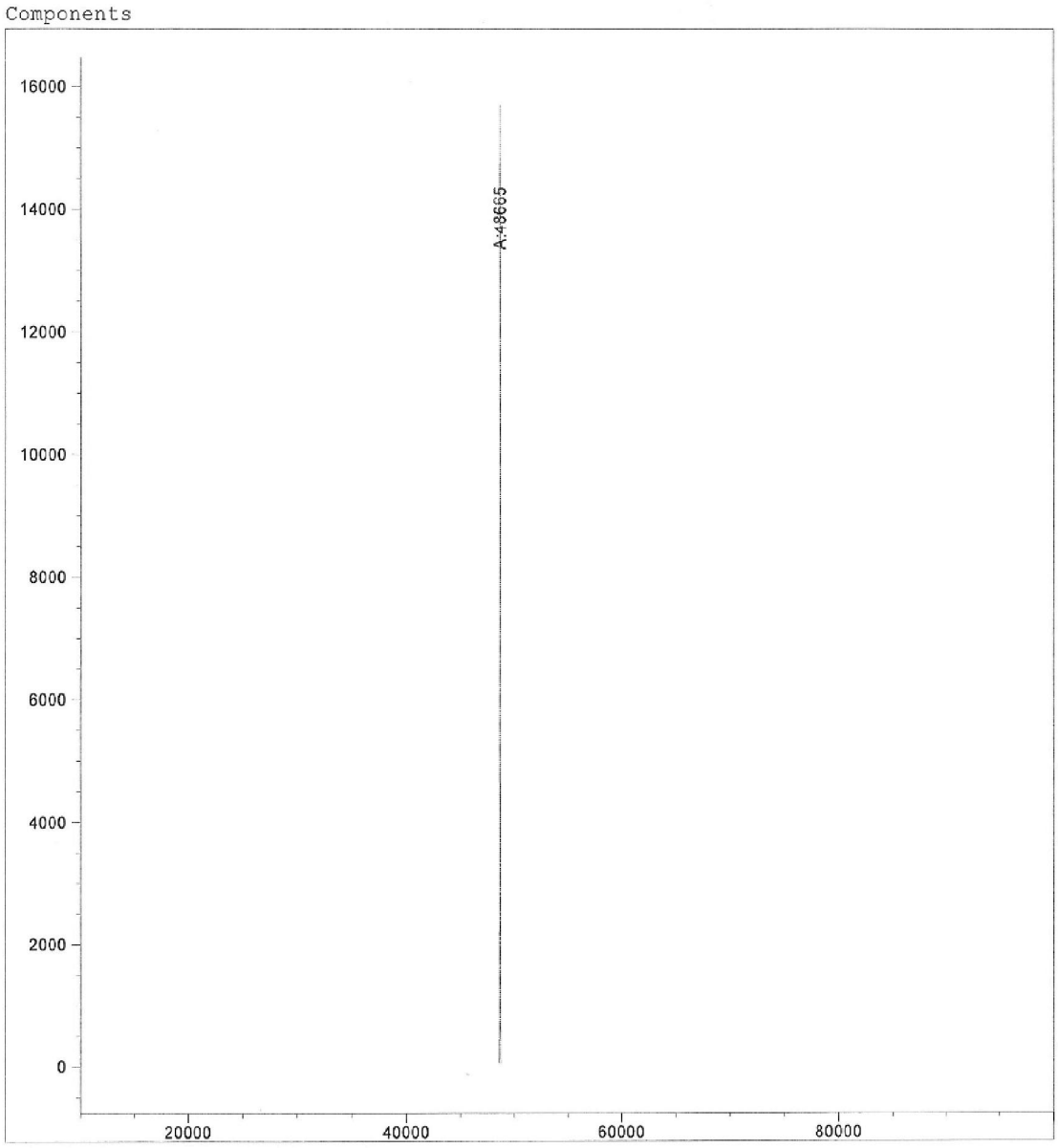

Fig. S5 (continued).

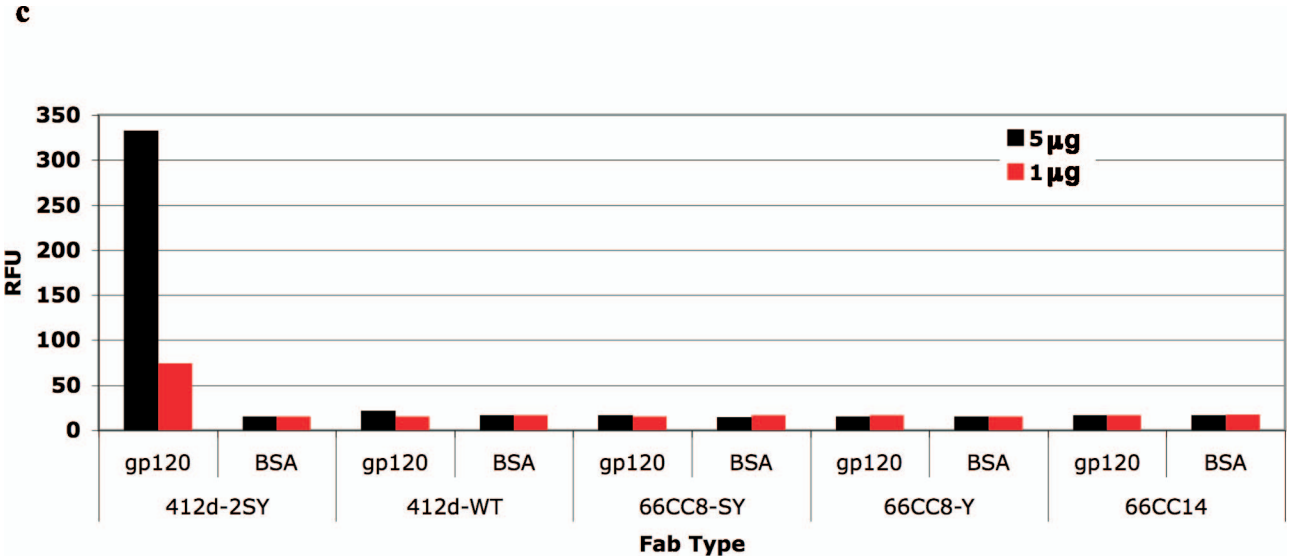

Fig. S5 (continued).

Table S1. Optimization of phage yield based on a model scFv

Synthetase plasmid backbone	Temperature (degrees °C)	Phage growth time (hr)	Strain used	Unnatural amino acid concentration (mM)	(Unnatural phage yield)/ (Natural phage yield)*
pSup	37	12	SY-X-*E. coli*	5	0.02
pSup	37	12	Keto-X-*E. coli*	5	0.02
pCDF	37	12	SY-X-*E. coli*	5	0.08
pCDF	37	12	Keto-X-*E. coli*	5	0.09
pCDF	30	12	Keto-X-*E. coli*	5	0.18
pCDF	30	18	Keto-X-*E. coli*	5	0.25
pCDF	30	18	SY-X-*E. coli*	10	0.24
pCDF	30	18	Keto-X-*E. coli*	8	0.40 ± 0.031
pCDF	30	18	SY-X-*E. coli*	15	0.33 ± 0.057
pCDF	30	18	Bpy-X-*E. coli*	1.5	0.34 ± 0.072
pCDF	30	18	Boro-X-*E. coli*	6.5	0.90 ± 0.149**

Strain used refers to the strain containing the unnatural amino acid and aminoacyl-tRNA synthetase specific to that strain. Unnatural amino acid was added directly to the media at the listed concentrations during phage expression. The yield ratio was taken as the titer of phage expressed from pSEX-GermTAG (unnatural phage) divided by the titer of phage expressed from pSEX-GermTAT (natural phage). We hypothesize that our optimization procedure increased the total production of unnatural amino acid containing pIII-fusion protein and at the same time decreased the rate of another step of the phage production process such as assembly or packaging.

* For the final set of conditions, ratios were determined in triplicate where cultures were split into three sets of two samples (natural and unnatural phage) for phage expression. (± standard deviations reported.)

** The addition of NaOH required to solubilize 4-borono-phenylalanine resulted in lower growth. Therefore, determination of natural phage yield was done in the presence of NaOH and 4-borono-phenylalanine in this case. In all other cases, addition of the unnatural amino acid did not noticeably affect natural phage yield.

Table S2. Final optimized conditions and yield bias in favor of phage expressing the test scFv-pIII with a TAT codon in place of the TAG codon (natural phage)

Strain	Synthetase plasmid backbone	Temperature (degrees °C)	Phage growth time (hr)	Unnatural amino acid concentration (mM)	Bias for natural phage
SY-X-*E. coli*	pCDF	30	18	15	3x
Keto-X-*E. coli*	pCDF	30	18	8	2.5x
Bpy-X-*E. coli*	pCDF	30	18	1.5	2.94x
Boro-X-*E. coli*	pCDF	30	18	6	1.11x

Table S3. Number of clones containing a TAG codon after phage expression from the pSEX-GermNNK library in Keto-X-*E. coli*, SY-X-*E. coli*, Bpy-X-*E. coli*, or Boro-X-*E. coli* (n = 50 or 100) and associated χ^2 values to show that bias on the population level is typified by bias of individual clone

Strain	# of clones sequenced (n)	# containing at least one TAG (measured)	# containing at least one TAG (expected from bias)	Expected standard deviation from binomial distribution*	χ^2 (using binomial distribution standard deviation)**
SY-X-*E. coli*	100	7	5.71	2.32	0.62
Keto-X-*E. coli*	100	10	6.92	1.79	2.94
Bpy-X-E. coli	100	7	5.88	1.66	1.62
Boro-X-*E. coli*	100	16	15.56	3.63	0.03
Before phage expression	50	9	8.65	2.67	0.03

Before phage expression entry refers to the clones found in the original library prior to phage production.

*$[np(1\text{-}p)]^{1/2}$ where n is # of clones sequenced and p = (unnatural phage yield) / (natural phage yield) $\times$ 0.173

**Values under 3.84 (corresponding to 5% probability) are accepted as consistent with the expected bias determined on an individual clone

Table S4. List of sequences selected from a 412d-based library doped with 412d-2SY

Sequence of doped 412d-2SY													
Y	P	N	D	*	N	D	*	A	P	E	E	G	M	
Selected sequences													
Y	T	N	D	L	N	D	*	G	E	E	E	H	G	
Y	N	N	D	D	N	D	L	R	L	G	E	*	S	
Y	D	N	D	N	D	D	G	T	A	E	E	*	Y	
Y	P	N	D	*	N	D	*	A	P	E	E	G	S	
Y	P	N	D	*	N	D	*	A	P	E	E	G	M	(~75% of population)

Underlined are the positions that were randomized with NNK.

Table S5. Enrichment for selection of gp120-binders from a 412d library

Round	Amount phage loaded	Amount phage eluted	Amount phage in final wash	Loaded/Eluted
1	1.23×10^{9}	4,000	<40	3.08×10^{5}
2	4.5×10^{8}	1.08×10^{4}	40	4.17×10^{4}
3	6.3×10^{6}	1.7×10^{4}	220	370
4	3×10^{7}	7,600	ND	3,747

Table S6. Percent phage containing TAG in 412d library selection rounds

Round	% phage containing at least one TAG	% phage containing the 412d2TAG doped clone
1	17.5 ($n = 40$)	2.2 ($n = 45$)
2	14.5 ($n = 55$)	10.9 ($n = 55$)
3	86 ($n = 50$)	74 ($n = 50$)
4	100 ($n = 20$)	100 ($n = 20$)

Table S7. Enrichment for selection of gp120-binders from a naive germline library

Round	Amount phage loaded	Amount phage eluted	Loaded/Eluted
1	1.28×10^{9}	2.0×10^{5}	6,400
2	2.72×10^{8}	9,200	2.86×10^{4}
3	1×10^{9}	2.4×10^{5}	4,166
4	2.5×10^{7}	3.5×10^{4}	714